A UNITARY EXTENSION PRINCIPLE FOR SHEARLET SYSTEMS

BIN HAN*, GITTA KUTYNIOK†, AND ZOUWEI SHEN‡

Abstract. In this paper, we first introduce the concept of an adaptive MRA (AMRA) structure which is a variant of the classical MRA structure suited to the main goal of a fast flexible decomposition strategy adapted to the data at each decomposition level. We then study this novel methodology for the general case of affine-like systems, and derive a Unitary Extension Principle (UEP) for filter design. Finally, we apply our results to the directional representation system of shearlets. This leads to a comprehensive theory for fast decomposition algorithms associated with shearlet systems which encompasses tight shearlet frames with spatially compactly supported generators within such an AMRA structure. Also shearlet-like systems associated with parabolic scaling and unimodular matrices optimally close to rotation as well as 3D shearlet systems are studied within this framework.

Key words. Affine systems, fast decomposition algorithm, shearlets, multiresolution analysis, tight frames

AMS subject classifications. Primary 42C40; Secondary 42C15, 65T60, 65T99, 94A08

1. Introduction. Wavelets are nowadays indispensable as a multiscale encoding system for a wide range of more theoretically to more practically oriented tasks, since they provide optimal approximation rates for smooth 1-dimensional data exhibiting singularities. The facts that they provide a unified treatment in both the continuous as well as digital setting and that the digital setting admits a multiresolution analysis leading to a fast spatial domain decomposition were essential for their success. It can however be shown that wavelets – although perfectly suited for isotropic structures – do not perform equally well when dealing with anisotropic phenomena.

This fact has motivated the development of various types of directional representation systems for 2-dimensional data that are capable of resolving edge- or curve-like features which precisely separate smooth regions in sparse way, of which examples are contourlets [19] and curvelets [10, 11, 12]. All these multiscale variants offer different advantages and disadvantages, however, neither of them provides a unified treatment of the continuous and digital setting. Curvelets, for instance, are known to yield tight frames but the digital curvelet transform is not designed within the curvelet-framework and hence, in particular, is not covered by the available theory [9].

About three years ago, a novel representation system – so-called shearlets – has been proposed [24, 38], which possesses the same favorable approximation and sparsity properties as the other candidates (see [25, 22, 39]) of whom curvelets are perhaps the most advanced ones. One main point in comparison with curvelets is the fact that angles are replaced by slopes when parameterizing directions which significantly supports the treating of the digital setting. A second main point is that shearlets fit within the general framework of affine-like systems, which provides an extensive

*Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada T6G 2G1 (bhan@math.ualberta.ca). B.H. is supported in part by NSERC Canada.

†Institute of Mathematics, University of Osnabrück, 49069 Osnabrück, Germany (kutyniok@math.uni-osnabrueck.de). G.K. would like to thank the Department of Mathematics at the National University of Singapore for its hospitality and support during her visit which enabled completion of a significant part of this paper.

‡Department of Mathematics, National University of Singapore, Singapore 117543 (matzuows@nus.edu.sg). Z.S. is supported in part by Grant R-146-000-113-112 at the National University of Singapore.

mathematical machinery. Thirdly, it is shown in [37] that shearlets – in addition to the aforementioned favorable properties – provide a unified treatment for the continuous and digital world similar to wavelets.

Recently, several researchers [26, 40, 41] have provided approaches to introduce an MRA structure with accompanying fast spatial domain decomposition for shearlets. This would establish shearlets as the directional representation system which provides the full range of advantageous properties for 2D data which wavelets provide in 1D. However, in the previous approaches, does either the MRA structure not lead to a tight frame, infinitely many filters make an implementation difficult, or the MRA structure is not faithful to the continuum transform. Approaches to extend the present shearlet theory to higher dimensions were also already undertaken, however, for now, only with continuous parameters [17].

In wavelet theory, the Unitary Extension Principle (UEP) introduced in [42] has proven to be a highly useful methodology for constructing tight wavelet frames with an associated MRA structure in order to develop efficient algorithms, for instance, for frame based image restorations including image deblurring and blind image deblurring, image inpainting, image denoising, and image decomposition (see [1, 2, 3, 4, 5, 6, 7, 8, 13, 14, 15, 16]). These applications of frame-based image restorations are also one motivation for our adventures here.

In this paper we aim at providing an MRA structure for tight shearlet frames – and in fact even for more general affine-like systems encompassing different shearlet systems as special cases – which exhibits all the favorable properties of MRA structures for wavelets. We also allow the MRA structure to be more flexible in the sense of adaptivity than ordinarily considered in the literature. We further prove sufficient conditions for such an adaptive MRA in terms of a suited Unitary Extension Principle (UEP) along with a fast spatial domain decomposition as well as approximation properties. Surprisingly, our general theory also includes shearlet systems for 3D data and provides fast decomposition algorithms for those.

1.1. List of Desiderata. The theoretical framework for a fast spatial domain decomposition we will develop in the case of general affine-like systems (cf. Subsection 1.2 for a precise definition), and, in particular, for shearlet systems shall satisfy the following desiderata:

(1) *Adaptivity.* The MRA structure shall allow interactive adaption of the decomposition procedure to different types of data either beforehand or during the process.
(2) *Computational Feasibility.* The MRA structure shall be suited to the fact that it aims for a decomposition algorithm, that means, for instance, that only a finite range of scales is required.
(3) *Fast Spatial Domain Decomposition.* An accompanying fast transform shall exist which decomposes digital data into high- and low-frequency parts, in particular, comprising directionality behavior.
(4) *Tight Frame Property.* This property will ensure stability of the decomposition and the possibility to employ the adjoint for reconstruction.
(5) *Compactly-Supportedness.* The framework shall allow the utilization of compactly supported generators to ensure spatial domain localization as well as fast computations, in particular, in the special case of shearlets.
(6) *Arbitrary Dimension.* The framework shall be applicable to affine-like systems for arbitrary dimension.

(7) *Faithfulness to Continuum Transform.* In the special case of shearlets, the transform shall be faithful to the continuum transform in the sense of being accompanied by a continuum shearlet system with discrete parameters.

1.2. An Adaptive Multiresolution Analysis (AMRA) for Affine-Like Systems. The framework of a multiresolution analysis is a well-established methodology in wavelet theory for deriving a decomposition of data into low- and high-frequency parts associated with a scaling function and wavelet which leads to a fast spatial domain decomposition. However, aiming for a decomposition with respect to general *affine-like systems*, which we coin systems being a subset of unions of *affine systems*

$$\bigcup_{\mathsf{M}\in\Lambda\subseteq GL(d,\mathbb{R})} \{\psi^\ell_{\mathsf{M}^j;\mathsf{k}} \ : \ j\in\mathbb{Z}, \mathsf{k}\in\mathbb{Z}^d, \ell=1,\dots,r\}, \qquad \psi^1,\dots,\psi^r\in L^2(\mathbb{R}^d),$$

where

$$\psi^\ell_{U;\mathsf{k}} := |\det U|^{1/2}\psi^\ell(U\cdot -\mathsf{k}), \qquad \mathsf{k}\in\mathbb{Z}^d, U\in GL(d,\mathbb{R}), \tag{1.1}$$

we are forced to reconsider this approach.

In light of desiderata (1)–(3) from the previous subsection and continuing the initial observations already made by one of the authors in [33], we claim a new paradigm for MRA manifested in the following two requirements to an MRA:

(R1) If one aims for a fast decomposition algorithm, it is sufficient to study non-homogeneous systems, i.e., considering a fixed finest decomposition level J instead of analyzing the limit $J\to\infty$. In the same way, the coarsest level might be defined to be the level 0 without loss of generality.

(R2) It is not necessary to specify beforehand which functions or masks generate the low- and which generate the high-frequency parts, since the key equation (Unitary Extension Principle) which needs to be satisfied at each decomposition level does not distinguish between them.

The main advantage of (R1) is to adapt the theoretical continuum considerations to the implementation requirements, whereas (R2) allows us to change the composition of the masks adaptively at each decomposition level. Therefore we coin this new paradigm for MRA an *Adaptive Multiresolution Analysis (AMRA).*

The decomposition technique will still rely on subdivision and transition operators which we in the sequel denote as follows: For a $d\times d$ invertible integer matrix M and a finitely supported sequence $a:\mathbb{Z}^d\to\mathbb{R}$, the *subdivision operator* $\mathrm{S}_{a,\mathsf{M}}: l(\mathbb{Z}^d)\to l(\mathbb{Z}^d)$ and the *transition operator* $\mathrm{T}_{a,\mathsf{M}}: l(\mathbb{Z}^d)\to l(\mathbb{Z}^d)$ are defined by

$$[\mathrm{S}_{a,\mathsf{M}}v](\mathsf{n}) := |\det(\mathsf{M})|\sum_{\mathsf{k}\in\mathbb{Z}^d} v(\mathsf{k})a(\mathsf{n}-\mathsf{M}\mathsf{k}) \quad \text{and} \quad [\mathrm{T}_{a,\mathsf{M}}v](\mathsf{n}) := \sum_{\mathsf{k}\in\mathbb{Z}^d} v(\mathsf{k})\overline{a(\mathsf{k}-\mathsf{M}\mathsf{n})}.$$

Note that $\widehat{\mathrm{S}_{a,\mathsf{M}}v}(\xi) = |\det(\mathsf{M})|\hat{v}(\mathsf{M}^T\xi)\hat{a}(\xi)$ and $\widehat{\mathrm{T}_{a,\mathsf{M}}v}(\mathsf{M}^T\xi) = |\det(\mathsf{M})|^{-1}\sum_{\omega\in\Omega_\mathsf{M}}\hat{v}(\xi+2\pi\omega)\overline{\hat{a}(\xi+2\pi\omega)}$, where $\hat{a}(\xi) := \sum_{\mathsf{k}\in\mathbb{Z}^d} a(\mathsf{k})e^{-i\mathsf{k}\cdot\xi}$ and $\Omega_\mathsf{M} := [(\mathsf{M}^T)^{-1}\mathbb{Z}^d]\cap[0,1)^d$.

We will now illustrate the general idea of an AMRA to the reader in the case of shearlet systems.

1.3. Shearlet Systems. Shearlet systems can be regarded as a special case of the previously introduced affine-like systems. The continuum shearlet transform with discrete parameters [24, 38] for functions in $L^2(\mathbb{R}^2)$ uses a two-parameter dilation

group, where one parameter indexes scale, and the second parameter indexes orientation. For each $c > 0$ and $s \in \mathbb{R}$, let A_c denote the *parabolic scaling matrix* and S_s denote the *shear matrix* of the form

$$A_c = \begin{pmatrix} c & 0 \\ 0 & \sqrt{c} \end{pmatrix} \qquad \text{and} \qquad S_s = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix},$$

respectively. To provide an equal treatment of the x- and y-axis, we split the frequency plane into the horizontal cone

$$\mathcal{C}_0 = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_1| \ge 1,\ |\xi_1/\xi_2| \ge 1\},$$

the vertical cone

$$\mathcal{C}_1 = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : |\xi_2| \ge 1,\ |\xi_1/\xi_2| \le 1\},$$

as well as a centered rectangle

$$\mathcal{R} = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \|(\xi_1, \xi_2)\|_\infty < 1\}$$

(see Figure 1.1 (a)).

(a) (b)

Fig. 1.1. *(a) The cones $\mathcal{C}_0$ and $\mathcal{C}_1$ and the centered rectangle $\mathcal{R}$ in frequency domain. (b) The tiling of the frequency domain induced by discrete shearlets.*

For cone $\mathcal{C}_1$, at scale $j \ge 0$, orientation $k = -2^j, \ldots, 2^j$, and spatial position $m \in \mathbb{Z}^2$, the associated *shearlets* are then defined by their Fourier transforms

$$\sigma_\eta = 2^{j\frac{3}{4}} \psi(S_k A_{4^j} \cdot - m),$$

where $\eta = (j, k, m, \iota)$ index scale, orientation, position, and cone. The shearlets for $\mathcal{C}_1$ are defined likewise by symmetry, as illustrated in Figure 1.1 (b), and we denote the resulting *shearlet system* by

$$\{\sigma_\eta : \eta \in \mathbb{N}_0 \times \{-2^j, \ldots, 2^j\} \times \mathbb{Z}^2 \times \{0, 1\}\}. \tag{1.2}$$

This is an affine-like system as defined before.

Notice that we chose a scaling of 4^j. Shearlet systems can be defined similarly for a scaling of 2^j, however, in this case the odd scales have to be handled particularly carefully.

One particular interesting shearlet generator is the function $\psi \in L^2(\mathbb{R}^2)$ defined by

$$\hat{\psi}(\xi) = \hat{\psi}(\xi_1, \xi_2) = \hat{\psi}_1(\xi_1)\, \hat{\psi}_2(\tfrac{\xi_2}{\xi_1}),$$

where $\psi_1 \in L^2(\mathbb{R})$ is a wavelet with $\hat{\psi}_1 \in C^\infty(\mathbb{R})$ and supp $\hat{\psi}_1 \subseteq [-4, -\frac{1}{4}] \cup [\frac{1}{4}, 4]$, and $\psi_2 \in L^2(\mathbb{R})$ is a 'bump' function satisfying $\hat{\psi}_2 \in C^\infty(\mathbb{R})$ and supp $\hat{\psi}_2 \subseteq [-1, 1]$. Filling in the low frequency band appropriately, with this particular generator the shearlet system (1.2) can be proven (see [24, Thm. 3]) to form a tight frame for $\{f \in L^2(\mathbb{R}^2) : \text{supp}\hat{f} \subset \mathcal{C}_0 \cup \mathcal{C}_1\}$. We remind the reader that a system X is called a *tight frame* for a Hilbert space $\mathcal{H}$ (in the literature sometimes also referred to as a *Parseval frame* or a *tight frame with bound one*), if $\|f\|^2 = \sum_{g\in X} |\langle f, g\rangle|^2$ holds for all $f \in \mathcal{H}$.

Concluding, the definition just discussed shows that shearlets live on anisotropic regions of width 2^{-2j} and length 2^{-j} at various orientations, which are parametrized by slope rather than angle as for second generation curvelets.

1.4. Shearlet AMRA. An intriguing application of the theory whose core ideas were presented in the Subsection 1.2 is the consideration of shearlet systems. In fact, this allows us to derive a framework which fulfills the desiderata we aimed for (see Subsection 1.1).

The fast spatial decomposition algorithm is derived as a special case of the methodology we will develop for general affine-like systems. To illustrate the idea of AMRA in the situation of shearlet systems, we will discuss one step of the decomposition algorithm. For this, let S_ℓ, $\ell = 1, \dots, r$ be a selection of 2×2 shear matrices, i.e., matrices of the type S_k, $k \in \mathbb{Z}$, let $1 \le s \le r$ be the separator between low- and high-frequency part, and let a_ℓ, $\ell = 1, \dots, r$ be finitely supported masks. Further, we denote the given data by v which for convenience purposes we now assume to lie in $l(\mathbb{Z}^2)$. Notice that certainly from the previous decomposition step we presumably have many such low-frequency coefficients. Retaining the notion introduced in Subsection 1.2, we then compute the next level of low-frequency coefficients by

$$v_\ell = \mathrm{T}_{a_\ell, S_\ell A_4} v, \qquad \ell = 1, \dots, s.$$

Similarly, we compute the next level of high-frequency coefficients by

$$v_\ell = \mathrm{T}_{a_\ell, S_\ell A_4} v, \qquad \ell = s+1, \dots, r.$$

The next step continues with decomposing v_ℓ, $\ell = 1, \dots, s$. The total number of decomposition steps is J, hence finite.

The reader should notice the requirements (R1) and (R2) from Subsection 1.2 as opposed to a 'classical MRA-decomposition algorithm'. (R1) is the finite number of decomposition steps, which is a more accurate model for what truly happens in a numerical decomposition. It thereby allows more flexibility, since the limit $J \to \infty$ doesn't need to be taken into account. (R2) can be seen by the fact that the separator between low- and high-frequency part is somehow 'loose' in the sense that at each level a certain condition needs to be satisfied (see Theorem 1.1 (i) below), which does *not* distinguish between those parts. This also implies that the structure of the subspaces the data is projected onto is not that strict than for a classical MRA, but allows also non-orthogonality and non-inclusiveness.

This algorithm is accompanied by a perfect reconstruction algorithm, if and only if, in each decomposition step the following version of the Unitary Extension Principle [42] is satisfied by the masks a_ℓ. For this, we have the following result, which we could coin the 'Shearlet Unitary Extension Principle'.

THEOREM 1.1. *Let S_ℓ, $\ell = 1, \dots, r$ be a selection of 2×2 shear matrices, i.e., matrices of the type S_k, $k \in \mathbb{Z}$, and let $\ell = 1, \dots, r$, be finitely supported masks. Then the following conditions are equivalent.*

(i) *For all* $v \in l(\mathbb{Z}^d)$,

$$\sum_{\ell=1}^{r} \mathrm{S}_{a_\ell, S_\ell A_4} \mathrm{T}_{a_\ell, S_\ell A_4} v = v.$$

(ii) *For any* $\omega \in \Omega = \bigcup_{\ell=1}^{r} \Omega_{S_\ell A_4}$, *where* $\Omega_{S_\ell A_4} := [(S_\ell^T)^{-1}(\frac{1}{4}\mathbb{Z} \times \frac{1}{2}\mathbb{Z})] \cap [0,1)^2$,

$$\sum_{\ell \in \{n \,:\, \omega \in \Omega_{S_n A_4}\}} \widehat{a_\ell}(\xi)\overline{\widehat{a_\ell}(\xi + 2\pi\omega)} = \delta(\omega),$$

where δ *denotes the Dirac sequence such that* $\delta(0) = 1$ *and* $\delta(\omega) = 0$ *for* $\omega \neq 0$.

In Section 2, we state a general version of this result (Theorem 2.1) from which this theorem can be derived as a corollary. A continuum version of it will then be discussed in Section 3. In Section 4, we will provide a variety of different choices for the masks, the dilation matrices, and the direction matrices. Adapted to the shearlet setting, the dilation matrices will typically be parabolic scaling matrices and the direction matrices will be chosen to be shear matrices.

1.5. Extensions and Further Questions. In fact, the results – including the general affine-like systems we claim to consider – are susceptible of extensive generalizations and extensions most of which are far beyond the scope of this paper and will be studied in future work.

- *Bi-Frame Case.* The bi-orthogonal framework for wavelets typically allows for significantly weakened filter conditions. Also for our more general MRA-framework, we can ask a similar question, which is then correctly stated as the bi-frame case.
- *Shearlets for Higher Dimensions.* Thriving applications such as the problem of geometric separation in image processing, for example, in astronomy when images of galaxies require separated analyses of stars, filaments, and sheets call for directional representation systems for 3D data, but also even higher dimensions (see, e.g., [20, 21]). Our framework conveys the potential to generate shearlets for higher dimensional signals alongside a fast flexible decomposition strategy, and we will present one example in 3D in Section 4.
- *Computational Realization.* Certainly, one extension of the work presented in this paper is the algorithmic study of the algorithms presented in this paper. Here we aim for understanding the additional flexibility these provide and how to optimally explore and utilize this property.

1.6. Related Work. Several research teams have previously designed MRA decomposition algorithms based on shearlets: we mention the affine system-based approach [26], the subdivision-based approach [40], and the approach based on separability [41]. However, neither of these approaches did satisfy all items of our list of desiderata (see Subsection 1.1). Further non-MRA based approaches were undertaken, for instance, in [23]. In our opinion, these pioneer efforts demonstrate real progress in directional representation, but further progress is needed to derive an in-all-aspects satisfactory comprehensive study of a fast spatial domain shearlet transform within an appropriate MRA framework with careful attention to mathematical exactness, faithfulness to the continuum transform, and computational feasibility, ideally fulfilling all our desiderata.

A particular credit deserves the work in [40], in which the adaptivity ideas were already lurking. The main difference to this paper is the additional freedom provided by the AMRA structure.

1.7. Contribution of this Paper. The contribution of this paper is three-fold. First, we introduce the concept of an adaptive MRA (AMRA) structure suited to the main goal of a fast flexible decomposition strategy. Secondly, we study this novel methodology for the general case of affine-like systems. And thirdly, we present a comprehensive theory for shearlet systems which encompasses tight shearlet frames with spatially compactly supported generators within such an AMRA structure along with a fast decomposition strategy.

1.8. Contents. In Section 2, we introduce the notation we employ for general affine-like systems, state the fast decomposition algorithm based on an AMRA, and prove the Unitary Extension Principle (UEP) for this situation. Section 3 is concerned with the relation to the continuum setting, i.e., with developing characterizing equations and approximation properties for the functions associated with an ARMA. This general methodology is then applied to the situation of shearlet systems in Section 4, where we present a general construction for associated filters. Here we also consider shearlet-like systems in the sense of integer-valued matrices approximating rotations different from the customarily employed shear matrices, and an approach of detecting anisotropic phenomena with an astonishingly simple isotropic system.

2. An Adaptive Multiresolution Analysis for General Affine-Like Systems. As already elaborated upon in the introduction, one main idea of an AMRA is to be able to design each decomposition step adaptively, for instance, dependent on the previous decomposition. To follow this philosophy, in Subsection 2.1, we will firstly analyze one single decomposition step, which might occur at any stage of the general decomposition algorithm. Secondly, in Subsection 2.2, we will then present the large picture in the sense of the complete decomposition procedure.

2.1. A Unitary Extension Principle for One Decomposition Step. Let now $v \in l(\mathbb{Z}^d)$ be some set of data. This could be the initial data, but also data after some steps of decomposition then on a renormalized grid. We assume that we are given a sequence of arbitrary $d \times d$ matrices M_ℓ, $1 \leq \ell \leq r$, and finitely supported masks a_ℓ, $\ell = 1, \ldots, r$ according to which the data shall be decomposed. Our first result is a Unitary Extension Principle (UEP) for this situation, which characterizes those matrices and filters, which allow perfect reconstruction from the decomposed data using subdivision.

THEOREM 2.1. *Let M_ℓ, $1 \leq \ell \leq r$ be $d \times d$ invertible integer matrices, and let a_ℓ, $\ell = 1, \ldots, r$, be finitely supported masks. Then the following conditions are equivalent.*

(i) *For all $v \in l(\mathbb{Z}^d)$,*

$$\sum_{\ell=1}^{r} \mathrm{S}_{a_\ell,\mathsf{M}_\ell} \mathrm{T}_{a_\ell,\mathsf{M}_\ell} v = v.$$

(ii) *For any $\omega \in \Omega = \bigcup_{\ell=1}^{r} \Omega_{\mathsf{M}_\ell}$, where $\Omega_{\mathsf{M}_\ell} := [(\mathsf{M}_\ell^T)^{-1}\mathbb{Z}^d] \cap [0,1)^d$,*

$$\sum_{\ell \in \{n \,:\, \omega \in \Omega_{\mathsf{M}_n}\}} \widehat{a_\ell}(\xi)\overline{\widehat{a_\ell}(\xi + 2\pi\omega)} = \delta(\omega). \tag{2.1}$$

Proof. First, notice that, for all $v \in l(\mathbb{Z}^d)$, (i) is equivalent to

$$\begin{aligned}
\hat{v}(\xi) &= \sum_{\ell=1}^{r} [\widehat{\mathrm{S}_{a_\ell,\mathsf{M}_\ell}\mathrm{T}_{a_\ell,\mathsf{M}_\ell}} v] \\
&= \sum_{\ell=1}^{r} |\det(\mathsf{M}_\ell)| \widehat{\mathrm{T}_{a_\ell,\mathsf{M}_\ell}} v(\mathsf{M}_\ell^T \xi)\widehat{a_\ell}(\xi) \\
&= \sum_{\ell=1}^{r} \sum_{\omega\in\Omega_{\mathsf{M}_\ell}} \hat{v}(\xi+2\pi\omega)\widehat{a_\ell}(\xi)\overline{\widehat{a_\ell}(\xi+2\pi\omega)}.
\end{aligned}$$

Employing the equivalence relation given by Ω_{M_n}, we can rewrite the previous equation as

$$\hat{v}(\xi) = \sum_{\omega\in\Omega} \sum_{\ell\in\{n\,:\,\omega\in\Omega_{\mathsf{M}_n}\}} \hat{v}(\xi+2\pi\omega)\widehat{a_\ell}(\xi)\overline{\widehat{a_\ell}(\xi+2\pi\omega)} \quad \text{for all} \quad v \in l(\mathbb{Z}^d). \tag{2.2}$$

Since (ii) implies (2.2), this proves (ii) $\Rightarrow$ (i).

We now assume that (i), hence (2.2) holds. We aim to prove that this implies (ii). For this, we let $B_\epsilon(\xi_0)$ denote as usual an open ball around ξ_0 with radius ϵ. We now first observe that for any $\omega_0 \in \Omega$ and any $\xi_0 \in \mathbb{R}$, there exists some $v \in l(\mathbb{Z}^d)$ and $\epsilon > 0$ such that

(a) $\hat{v}(\xi+2\pi\omega_0) = 1$ for all $\xi \in B_\epsilon(\xi_0)$,

(b) $\hat{v}(\xi+2\pi\omega) = 0$ for all $\xi \in B_\epsilon(\xi_0), \omega \in \Omega \setminus \{\omega_0\}$,

(c) $\operatorname{supp}\hat{v} \subseteq 2\pi\omega_0 + B_{2\epsilon}(\xi_0)$,

simply since Ω is discrete. Now we can conclude that (2.2) implies

$$\hat{v}(\xi) = \sum_{\ell\in\{n\,:\,\omega\in\Omega_{\mathsf{M}_n}\}} \widehat{a_\ell}(\xi)\overline{\widehat{a_\ell}(\xi+2\pi\omega)} \quad \text{for all } \omega\in\Omega,\ \xi\in B_\epsilon(\xi_0).$$

Hence again by (a) – (c), condition (ii) follows for all $\xi \in B_\epsilon(\xi_0)$. Since ξ_0 is arbitrary chosen, condition (ii) follows. ∎

The first part (equivalence of conditions (i) and (ii)) of the next result follows immediately from Theorem 2.1 as a special case. It significantly simplifies the condition on the mask imposed by the UEP, provided the situation allows to always use the same sampling lattice. Condition (iii) is the spatial domain expression for the UEP condition (ii), which illustrates the filter design problem in spatial domain. Certainly, this condition could also be stated in the general situation of Theorem 2.1. We however decide to omit this, since the content would be clouded by very technical details.

Corollary 2.2. *Let M_ℓ, $1 \le \ell \le r$ be $d\times d$ invertible integer matrices satisfying that $\mathsf{M}_\ell\mathbb{Z}^d = \mathsf{M}\mathbb{Z}^d$ for all $1 \le \ell \le r$ for some matrix M, and let a_ℓ, $\ell = 1,\ldots,r$, be finitely supported masks. Then the following conditions are equivalent.*

(i) *For all $v \in l(\mathbb{Z}^d)$,*

$$\sum_{\ell=1}^{r} \mathrm{S}_{a_\ell,\mathsf{M}_\ell}\mathrm{T}_{a_\ell,\mathsf{M}_\ell} v = v.$$

(ii) *For any $\omega \in \Omega_{\mathsf{M}} = [(\mathsf{M}^T)^{-1}\mathbb{Z}^d] \cap [0,1)^d$,*

$$\sum_{\ell=1}^{r} \widehat{a_\ell}(\xi)\overline{\widehat{a_\ell}(\xi+2\pi\omega)} = \delta(\omega).$$

(iii) *For all* $\mathsf{k}, \gamma \in \mathbb{Z}^d$,

$$\sum_{\mathsf{n}\in\mathbb{Z}^2} \overline{a_\ell(\mathsf{k}+\mathsf{M}\mathsf{n}+\gamma)}a_\ell(\mathsf{M}\mathsf{n}+\gamma) = |\det(\mathsf{M})|^{-1}\delta(\mathsf{k}).$$

Proof. As already mentioned, the equivalence of (i) and (ii) does follow from Theorem 2.1.

We next prove equivalence between (ii) and (iii). By using the definition of $\widehat{a_\ell}$, condition (ii) is equivalent to

$$\begin{aligned}\delta(\omega) &= \sum_{\ell=1}^{r}\sum_{\mathsf{k}\in\mathbb{Z}^d}\overline{a_\ell(\mathsf{k})}e^{i\mathsf{k}\cdot\xi}\sum_{\mathsf{n}\in\mathbb{Z}^d}a_\ell(\mathsf{n})e^{-i\mathsf{n}\cdot(\xi+2\pi\omega)}\\ &= \sum_{\ell=1}^{r}\sum_{\mathsf{k},\mathsf{n}\in\mathbb{Z}^d}\overline{a_\ell(\mathsf{k})}a_\ell(\mathsf{n})e^{i(\mathsf{k}-\mathsf{n})\cdot\xi}e^{-i\mathsf{n}\cdot 2\pi\omega}. \end{aligned}\tag{2.3}$$

Next we denote $\Gamma_{\mathsf{M}} = \mathbb{Z}^d/[\mathsf{M}\mathbb{Z}^d]$ with $0 \in \Gamma_{\mathsf{M}}$. Then

$$\mathbb{Z}^d = \Gamma_{\mathsf{M}} + \mathsf{M}\mathbb{Z}^d$$

follows directly. This now allows us to rewrite (2.3) as

$$\begin{aligned}\delta(\omega) &= \sum_{\ell=1}^{r}\sum_{\gamma\in\Gamma_{\mathsf{M}}}\sum_{\mathsf{k},\mathsf{n}\in\mathbb{Z}^d}\overline{a_\ell(\mathsf{k})}a_\ell(\mathsf{M}\mathsf{n}+\gamma)e^{i(\mathsf{k}-\mathsf{M}\mathsf{n}-\gamma)\cdot\xi}e^{-i(\mathsf{M}\mathsf{n}+\gamma)2\pi\omega}\\ &= \sum_{\ell=1}^{r}\sum_{\gamma\in\Gamma_{\mathsf{M}}}\sum_{\mathsf{k},\mathsf{n}\in\mathbb{Z}^d}\overline{a_\ell(\mathsf{k}+\mathsf{M}\mathsf{n}+\gamma)}a_\ell(\mathsf{M}\mathsf{n}+\gamma)e^{i\mathsf{k}\cdot\xi}e^{-i\gamma\cdot 2\pi\omega}.\end{aligned}$$

Considering this equation in matrix form

$$\left(e^{-i\gamma\cdot 2\pi\omega}\right)_{\omega\in\Omega_{\mathsf{M}},\gamma\in\Gamma_{\mathsf{M}}}\left(\sum_{\mathsf{k},\mathsf{n}\in\mathbb{Z}^d}\overline{a_\ell(\mathsf{k}+\mathsf{M}\mathsf{n}+\gamma)}a_\ell(\mathsf{M}\mathsf{n}+\gamma)e^{i\mathsf{k}\cdot\xi}\right)_{\gamma\in\Gamma_{\mathsf{M}}} = \left(\delta(\omega)\right)_{\omega\in\Omega_{\mathsf{M}}},$$

we can conclude by taking the inverse that

$$\begin{aligned}&\left(\sum_{\mathsf{k},\mathsf{n}\in\mathbb{Z}^d}\overline{a_\ell(\mathsf{k}+\mathsf{M}\mathsf{n}+\gamma)}a_\ell(\mathsf{M}\mathsf{n}+\gamma)e^{i\mathsf{k}\cdot\xi}\right)_{\gamma\in\Gamma_{\mathsf{M}}}\\ &= |\det(\mathsf{M}_1)|^{-1}\left(e^{i\gamma\cdot 2\pi\omega}\right)_{\omega\in\Omega_{\mathsf{M}},\gamma\in\Gamma_{\mathsf{M}}}\left(\delta(\omega)\right)_{\omega\in\Omega_{\mathsf{M}}}\\ &= |\det(\mathsf{M})|^{-1}\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}.\end{aligned}$$

Thus (ii) is equivalent to the equation

$$\sum_{\mathsf{k},\mathsf{n}\in\mathbb{Z}^d}\overline{a_\ell(\mathsf{k}+\mathsf{M}\mathsf{n}+\gamma)}a_\ell(\mathsf{M}\mathsf{n}+\gamma)e^{i\mathsf{k}\cdot\xi} = |\det(\mathsf{M})|^{-1} \quad \text{for all } \gamma\in\mathbb{Z}^d,$$

which in turn is equivalent to (iii). ∎

This corollary implies, by defining the equivalence relation $\sim$ on the $d \times d$ invertible integer matrices by

$$\mathsf{M} \sim \mathsf{M}' \quad :\Longleftrightarrow \quad \mathsf{M}\mathbb{Z}^d = \mathsf{M}'\mathbb{Z}^d,$$

that only one representative of each involved class needs to satisfy (2.1). In other words, only the generated lattices $\mathsf{M}\mathbb{Z}^d$ do matter in the condition equivalent to perfect reconstruction. Inside each class we have very much freedom to choose the dilation matrices as necessary by the application. We refer the reader to [29] for the application of this observation on the construction of wavelets.

2.2. Fast Decomposition Algorithm. In the previous subsection, we derived characterizing conditions for sequences of arbitrary $d \times d$ matrices M_ℓ, $1 \le \ell \le r$, and finitely supported masks a_ℓ, $\ell = 1, \ldots, r$ which allow perfect reconstruction in each step of an AMRA. This now enables us to present a general adaptive decomposition algorithm, where the matrices and masks in each step can be chosen according to Theorem 2.1.

We first require some notation to carefully keep track of the decomposition steps and positions in the generated tree structure by suitable indexes. While reading the definitions, we recommend the reader to also take a look at Figure 2.1, which illustrates the indexing of the decomposition. The heart of our indices are vectors $(\beta_1, \ldots, \beta_J)$ assigned to matrices, filters, and data, where each entry β_j indicates whether this object is related to the computation of scale j or a coarser scale, and if yes, whether the data at scale j to reach this object was generated by a low-pass or high-pass filter.

Let us now be more specific. The original data is assigned the index $\mathbf{0} := (0, \ldots, 0) \in \mathbb{N}_0^J$, and we set $\mathcal{L}_0^L = \{\mathbf{0}\} = \{(0, \ldots, 0)\}$. For the indexing of the low-pass filters, we define $\mathcal{I}_{\mathbf{0}}^L = \{1, \ldots, s_{\mathbf{0}}\}$ and, for the high-pass filters, $\mathcal{I}_{\mathbf{0}}^H = \{s_{\mathbf{0}}+1, \ldots, r_{\mathbf{0}}\}$, where $s_{\mathbf{0}} \le r_{\mathbf{0}}$ and $r_{\mathbf{0}}$ is a positive integer. In that sense $s_{\mathbf{0}}$ *partitions* the filters into low- and high-frequency filters. If $s_{\mathbf{0}} = r_{\mathbf{0}}$, then $\mathcal{I}_{\mathbf{0}}^H = \emptyset$. Thus, matrices and filters in the first step of the decomposition are labeled by $(\ell, 0, \ldots, 0) \in \mathbb{N}_0^J$, $\ell \in \mathcal{I}_{\mathbf{0}}^L$ for the low-frequency objects, and $(\ell, 0, \ldots, 0) \in \mathbb{N}_0^J$, $\ell \in \mathcal{I}_{\mathbf{0}}^H$ for the high-frequency objects. For our convenience, we further introduce the notation

$$\mathcal{L}_1^L = \mathcal{L}_{1,\mathbf{0}}^L = \{(\ell, 0, \ldots, 0) \in \mathbb{R}^J : \ell \in \mathcal{I}_{\mathbf{0}}^L\}$$

and

$$\mathcal{L}_1^H = \mathcal{L}_{1,\mathbf{0}}^H = \{(\ell, 0, \ldots, 0) \in \mathbb{R}^J : \ell \in \mathcal{I}_{\mathbf{0}}^H\}$$

to denote the indices for the low- and high-frequency objects in the first decomposition layer. The associated filters $a_\beta, \beta \in \mathcal{L}_1^L \cup \mathcal{L}_1^H$ for the first decomposition step are required to satisfy condition (ii) in Theorem 2.1, i.e.,

$$\sum_{\beta \in \{\gamma \in \mathcal{L}_1^L \cup \mathcal{L}_1^H \,:\, \omega \in \Omega_{\mathsf{M}_\gamma}\}} \widehat{a_\beta}(\xi)\overline{\widehat{a_\beta}(\xi + 2\pi\omega)} = \delta(\omega), \quad \omega \in \Omega_{\mathbf{0}} = \Omega_{(0,\ldots,0)},$$

where M_γ are $d \times d$ invertible integer matrices, $\Omega_{\mathsf{M}_\gamma} := [(\mathsf{M}_\gamma^T)^{-1}\mathbb{Z}^d] \cap [0,1)^d$, and $\Omega_{(0,\ldots,0)} := \cup_{\gamma \in \mathcal{L}_1^L \cup \mathcal{L}_1^H} \Omega_{\mathsf{M}_\gamma}$.

The matrices and filters which are used in the jth step will then be labeled as follows. We first assume that in the $(j-1)$th step the sets $\mathcal{L}_{j-1}^L$ and $\mathcal{L}_{j-1}^H$ were already constructed. Then, for some $\beta = (\beta_1, \ldots, \beta_{j-1}, 0, \ldots, 0) \in \mathcal{L}_{j-1}^L$, we define

$$\mathcal{I}_\beta^L = \{1, \ldots, s_\beta\} \quad \text{and} \quad \mathcal{I}_\beta^H = \{s_\beta + 1, \ldots, r_\beta\} \quad \text{with } 1 \le s_\beta \le r_\beta,$$

as single labels in the new decomposition layer for the matrices and filters. To take the whole tree structure into account, we now define the set of new low-pass related indices arising from $\beta = (\beta_1, \dots, \beta_{j-1}, 0, \dots, 0)$ by

$$\mathcal{L}^L_{j,\beta} = \{(\beta_1, \dots, \beta_{j-1}, \ell, 0, \dots, 0) \in \mathbb{N}_0^J : \ell \in \mathcal{I}^L_\beta\}$$

and the same for the set of all high-pass related indices,

$$\mathcal{L}^H_{j,\beta} = \{(\beta_1, \dots, \beta_{j-1}, \ell, 0, \dots, 0) \in \mathbb{N}_0^J : \ell \in \mathcal{I}^H_\beta\}.$$

We further set

$$\mathcal{L}^L_j = \cup_{\beta \in \mathcal{L}^L_{j-1}} \mathcal{L}^L_{j,\beta} \quad \text{and} \quad \mathcal{L}^H_j = \cup_{\beta \in \mathcal{L}^L_{j-1}} \mathcal{L}^H_{j,\beta},$$

the complete set of low-pass and high-pass indices in step j, respectively. Now let $\alpha \in \mathcal{L}^L_{j-1}$. To ensure perfect reconstruction, according to Theorem 2.1, the associated filters a_β, $\beta \in \mathcal{L}^L_{j,\alpha} \cup \mathcal{L}^H_{j,\alpha}$ must satisfy

$$\sum_{\beta \in \{\gamma \in \mathcal{L}^L_{j,\alpha} \cup \mathcal{L}^H_{j,\alpha} \,:\, \omega \in \Omega_{\mathsf{M}_\gamma}\}} \widehat{a_\beta}(\xi)\overline{\widehat{a_\beta}(\xi + 2\pi\omega)} = \delta(\omega), \qquad \omega \in \Omega_\alpha,$$

where $\Omega_\alpha := \cup_{\gamma \in \mathcal{L}^L_{j,\alpha} \cup \mathcal{L}^H_{j,\alpha}} \Omega_{\mathsf{M}_\gamma}$.

Next we describe the general multi-level decomposition algorithm explicitly using the indexing that we just introduced. For illustrative purposes, before presenting the complete algorithm, we display the first decomposition step as well as one part of the second decomposition step in Figure 2.1.

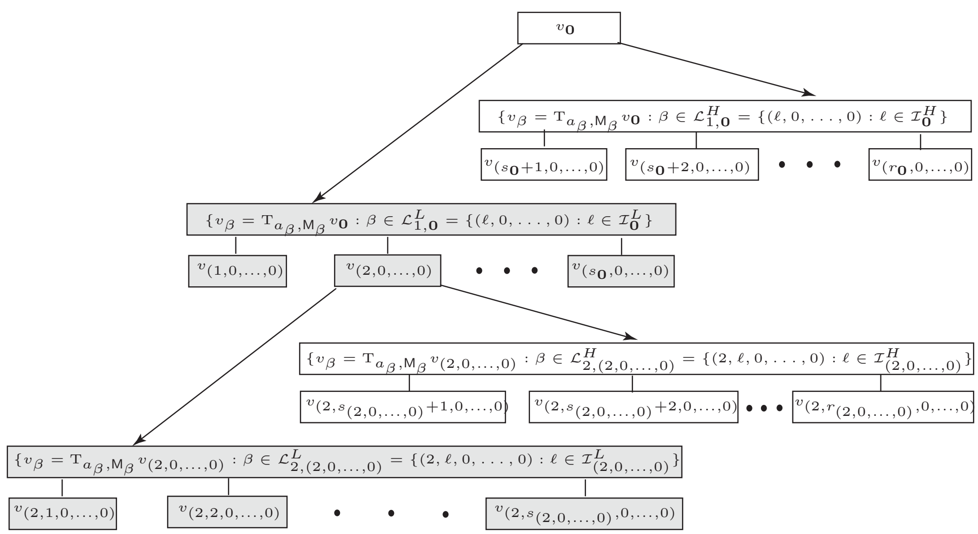

FIG. 2.1. *The decomposition structure of* (FAD) *illustrated through the first complete step and the second step shown exemplarily through the decomposition of* $v_{(2,0,\dots,0)}$. *The low-frequency components, which will be processed further, are gray-shaded.*

In Figure 2.2, we now describe the general multi-level decomposition algorithm explicitly. This decomposition can be implemented as the usual fast wavelet transform with a tree structure. We remind the reader of the introduced notion of transition and subdivision in Subsection 1.2.

FAD (FAST ADAPTIVE DECOMPOSITION)

Parameters:

- Data $v \in l(\mathbb{Z}^d)$.
- Number of iterations $J \in \mathbb{N}$.
- Set of all indices $\mathcal{L} := \bigcup_{j=1}^{J} \mathcal{L}_j^L \cup \bigcup_{j=1}^{J} \mathcal{L}_j^H$.
- Number of decompositions r_β for each position $\beta \in \mathcal{L}_{j-1}^L$, $j = 1, \ldots, J$.
- Separators between low- and high-pass parts $s_\beta \le r_\beta$ for each position $\beta \in \mathcal{L}_{j-1}^L$, $j = 1, \ldots, J$.
- Sequence of finitely supported masks $a_\beta \in l(\mathbb{Z}^2)$, $\beta \in \mathcal{L}$, satisfying the conditions in Theorem 2.1.
- Sequence of matrices $(\mathsf{M}_\beta)_{\beta \in \mathcal{L}}$ satisfying the conditions in Theorem 2.1.

Algorithm:

Set $v_{\mathbf{0}} = v_{(0,\ldots,0)} = v$.
For $j = 1, \ldots, J$ do
    For $\beta \in \mathcal{L}_{j,\alpha}^L$, $\alpha \in \mathcal{L}_{j-1}^L$ do
        Set $v_\beta = \mathrm{T}_{a_\beta, \mathsf{M}_\beta} v_\alpha$.
    end;
    For $\beta \in \mathcal{L}_{j,\alpha}^H$, $\alpha \in \mathcal{L}_{j-1}^L$ do
        Set $w_\beta = \mathrm{T}_{a_\beta, \mathsf{M}_\beta} v_\alpha$.
    end;
end;

Output:

- Low-pass coefficients v_β, where $\beta \in \mathcal{L}_J^L$.
- High-pass coefficients w_β, where $\beta \in \bigcup_{j=1}^{J} \mathcal{L}_j^H$.

FIG. 2.2. *The* FAD *Algorithm for a fast adaptive decomposition using affine-like systems.*

The fact that the filters are chosen to be perfect reconstruction filters allows us to reconstruct the data by application of appropriate subdivision operators as displayed in Figure 2.3.

3. Relation with Continuum Setting.

3. Relation with Continuum Setting. We next intend to relate the 'digital conditions' on the filters to 'continuum conditions' for associated function systems, in particular, frame systems.

3.1. Characterization Equations. As before in Theorem 2.1, we only consider one level. The conditions we will derive then need to be satisfied for each step in the iteration, and the choice of a non-stationary or stationary scheme is left to the user.

For a function ϕ and invertible $d \times d$ matrix U, we shall use the notation as in (1.1).

We can now formulate the condition on prefect reconstruction in terms of a condition in the function setting. Notice that here we again consider the most general situation of Theorem 2.1 as opposed to Corollary 2.2.

THEOREM 3.1. *Let M_ℓ, $0 \le \ell \le r$ be $d \times d$ invertible integer matrices, and let a_ℓ, $\ell = 1, \ldots, r$, be finitely supported masks. Let ϕ be a nontrivial compactly supported function in $L^2(\mathbb{R}^d)$. Then the following conditions are equivalent.*

FAR (Fast Adaptive Reconstruction)

Parameters:
- Number of iterations $J \in \mathbb{N}$.
- Set of all indices $\mathcal{L} := \bigcup_{j=1}^{J} \mathcal{L}_j^L \cup \bigcup_{j=1}^{J} \mathcal{L}_j^H$.
- Number of reconstructions r_β for each position $\beta \in \mathcal{L}_{j-1}^L$, $j = 1, \ldots, J$.
- Separators between low- and high-pass parts $s_\beta \leq r_\beta$ for each position $\beta \in \mathcal{L}_{j-1}^L$, $j = 1, \ldots, J$.
- Low-pass coefficients v_β, where $\beta \in \mathcal{L}_J^L$.
- High-pass coefficients w_β, where $\beta \in \bigcup_{j=1}^{J} \mathcal{L}_j^H$.
- Sequence of finitely supported masks $a_\beta \in l(\mathbb{Z}^2)$, $\beta \in \mathcal{L}$, satisfying the conditions in Theorem 2.1.
- Sequence of matrices $(\mathsf{M}_\beta)_{\beta \in \mathcal{L}}$ satisfying the conditions in Theorem 2.1.

Recursive Algorithm:

1) For $j = J-1, \ldots, 0$ do
2) For $\alpha \in \mathcal{L}_j^L$ do
3) Set $v_\alpha = \sum_{\beta \in \mathcal{L}_{j+1,\alpha}^L} \mathrm{S}_{a_\beta, \mathsf{M}_\beta} v_\beta + \sum_{\gamma \in \mathcal{L}_{j+1,\alpha}^H} \mathrm{S}_{a_\gamma, \mathsf{M}_\gamma} w_\gamma$.
4) end;
5) end;

Output:
- Data $v_{\mathbf{0}} = v_{(0,\ldots,0)} \in l(\mathbb{Z}^d)$.

Fig. 2.3. *The* FAR *Algorithm for a fast adaptive reconstruction using affine-like systems.*

(i) *For all* $v \in l(\mathbb{Z}^d)$,

$$\sum_{\ell=1}^{r} \mathrm{S}_{a_\ell, \mathsf{M}_\ell} \mathrm{T}_{a_\ell, \mathsf{M}_\ell} v = v.$$

(ii) *For each* $f, g \in L^2(\mathbb{R}^d)$,

$$\sum_{\mathsf{k} \in \mathbb{Z}^d} \langle f, \phi_{\mathsf{M}_0;\mathsf{k}} \rangle \langle \phi_{\mathsf{M}_0;\mathsf{k}}, g \rangle = \sum_{\ell=1}^{r} \sum_{\mathsf{k} \in \mathbb{Z}^d} \langle f, \psi^\ell_{\mathsf{M}_\ell^{-1}\mathsf{M}_0;\mathsf{k}} \rangle \langle \psi^\ell_{\mathsf{M}_\ell^{-1}\mathsf{M}_0;\mathsf{k}}, g \rangle, \tag{3.1}$$

where $\phi_{\mathsf{M}_0;\mathsf{k}}$ *is defined as in* (1.1) *and* $\psi^1, \ldots, \psi^r$ *are defined by* $\widehat{\psi^\ell}(\mathsf{M}_\ell^T \xi) := \widehat{a_\ell}(\xi)\hat{\phi}(\xi)$, *that is,*

$$\psi^\ell := |\det(\mathsf{M}_\ell)| \sum_{\mathsf{k} \in \mathbb{Z}^d} a_\ell(\mathsf{k}) \phi(\mathsf{M}_\ell \cdot - \mathsf{k}).$$

Proof. We first recall that (i) is equivalent to (2.1) by Theorem 2.1. The strategy will now be to show that (ii) is equivalent to (2.1).

We first assume that (2.1) holds. Then $\langle f, \phi_{\mathsf{M}_0;\mathsf{k}} \rangle = \langle |\det \mathsf{M}_0|^{-1/2} f(\mathsf{M}_0^{-1} \cdot), \phi(\cdot - \mathsf{k}) \rangle = \langle |\det \mathsf{M}_0|^{-1/2} f(\mathsf{M}_0^{-1} \cdot), \phi_{I_d;\mathsf{k}} \rangle$. It is easy to see that (3.1) holds if and only if it holds for $\mathsf{M}_0 = I_d$. So, we assume that $\mathsf{M}_0 = I_d$ in the following proof. Using the Fourier-based approach as in [33, Lemma 3], we have

$$\sum_{\mathsf{k} \in \mathbb{Z}^d} \langle f, \phi_{I_d;\mathsf{k}} \rangle \langle \phi_{I_d;\mathsf{k}}, g \rangle = (2\pi)^d \int_{\mathbb{R}^d} \sum_{\mathsf{k} \in \mathbb{Z}^d} \hat{f}(\xi) \overline{\hat{g}(\xi + 2\pi\mathsf{k})}\, \overline{\hat{\phi}(\xi)} \hat{\phi}(\xi + 2\pi\mathsf{k}) d\xi$$

and

$$\begin{aligned}
&\sum_{\ell=1}^{r}\sum_{\mathsf{k}\in\mathbb{Z}^d}\langle f,\psi^{\ell}_{\mathsf{M}_\ell^{-1};\mathsf{k}}\rangle\langle\psi^{\ell}_{\mathsf{M}_\ell^{-1};\mathsf{k}},g\rangle\\
&=(2\pi)^d\int_{\mathbb{R}^d}\sum_{\ell=1}^{r}\sum_{\mathsf{k}\in\mathbb{Z}^d}\hat f(\xi)\overline{\hat g(\xi+2\pi(\mathsf{M}_\ell^T)^{-1}\mathsf{k})}\,\overline{\widehat{\psi^\ell}(\mathsf{M}_\ell^T\xi)}\widehat{\psi^\ell}(\mathsf{M}_\ell^T\xi+2\pi\mathsf{k})d\xi.
\end{aligned}\tag{3.2}$$

Since $\widehat{\psi^\ell}(\mathsf{M}_\ell^T\xi):=\widehat{a_\ell}(\xi)\hat\phi(\xi)$, we have

$$\overline{\widehat{\psi^\ell}(\mathsf{M}_\ell^T\xi)}\widehat{\psi^\ell}(\mathsf{M}_\ell^T\xi+2\pi\mathsf{k})=\overline{\widehat{a_\ell}(\xi)}\widehat{a_\ell}(\xi+2\pi(\mathsf{M}_\ell^T)^{-1}\mathsf{k})\overline{\hat\phi(\xi)}\hat\phi(\xi+2\pi(\mathsf{M}_\ell^T)^{-1}\mathsf{k}).$$

Note that $\mathbb{Z}^d$ is the disjoint union of $\mathsf{M}_\ell^T\omega+\mathsf{M}_\ell^T\mathbb{Z}^d$, $\omega\in\Omega_{\mathsf{M}_\ell}$. Now we deduce that (3.2) becomes

$$\begin{aligned}
&\sum_{\ell=1}^{r}\sum_{\mathsf{k}\in\mathbb{Z}^d}\langle f,\psi^{\ell}_{\mathsf{M}_\ell^{-1};\mathsf{k}}\rangle\langle\psi^{\ell}_{\mathsf{M}_\ell^{-1};\mathsf{k}},g\rangle\\
&=(2\pi)^d\int_{\mathbb{R}^d}\sum_{\ell=1}^{r}\sum_{\omega\in\Omega_{\mathsf{M}_\ell}}\sum_{\mathsf{k}\in\mathbb{Z}^d}\hat f(\xi)\overline{\hat g(\xi+2\pi\omega+2\pi\mathsf{k})}\,\overline{\widehat{a_\ell}(\xi)}\widehat{a_\ell}(\xi+2\pi\omega)\\
&\cdot\overline{\hat\phi(\xi)}\hat\phi(\xi+2\pi\omega+2\pi\mathsf{k})d\xi\\
&=(2\pi)^d\int_{\mathbb{R}^d}\sum_{\omega\in\Omega}\sum_{\mathsf{k}\in\mathbb{Z}^d}\hat f(\xi)\overline{\hat g(\xi+2\pi\omega+2\pi\mathsf{k})}\,\overline{\hat\phi(\xi)}\hat\phi(\xi+2\pi\omega+2\pi\mathsf{k})\\
&\cdot\sum_{\ell\in\{1\le n\le r\,:\,\omega\in\Omega_{\mathsf{M}_n}\}}\overline{\widehat{a_\ell}(\xi)}\widehat{a_\ell}(\xi+2\pi\omega)d\xi.
\end{aligned}$$

Hence, (3.1) is equivalent to

$$\begin{aligned}
&\int_{\mathbb{R}^d}\sum_{\mathsf{k}\in\mathbb{Z}^d}\hat f(\xi)\overline{\hat g(\xi+2\pi\mathsf{k})}\,\overline{\hat\phi(\xi)}\hat\phi(\xi+2\pi\mathsf{k})d\xi\\
&=\int_{\mathbb{R}^d}\sum_{\omega\in\Omega}\sum_{\mathsf{k}\in\mathbb{Z}^d}\hat f(\xi)\overline{\hat g(\xi+2\pi\omega+2\pi\mathsf{k})}\,\overline{\hat\phi(\xi)}\hat\phi(\xi+2\pi\omega+2\pi\mathsf{k})\\
&\quad\cdot\sum_{\ell\in\{1\le n\le r\,:\,\omega\in\Omega_{\mathsf{M}_n}\}}\overline{\widehat{a_\ell}(\xi)}\widehat{a_\ell}(\xi+2\pi\omega)d\xi.
\end{aligned}\tag{3.3}$$

Since (2.1) holds, it is obvious that the above identity in (3.3) holds and therefore, (iii) holds.

Conversely, if (ii) holds, then (3.3) holds. Now we use a similar argument as in [33, Lemma 5] to prove that this implies condition (ii). For this, we note that $\Omega+\mathbb{Z}^d:=\{\omega+\mathsf{k}\,:\,\omega\in\Omega,\mathsf{k}\in\mathbb{Z}^d\}$ is a discrete set without any accumulation point. For any $\omega_0\in\Omega$ and $\xi_0\in\mathbb{R}^d$, we can take any functions $f,g\in L^2(\mathbb{R}^d)$ such that the support of $\hat f$ is contained inside $B_\epsilon(\xi_0)$ and the support of $\hat g$ is contained inside $B_\epsilon(\xi+2\pi\omega_0)$. As long as ϵ is small enough, it is not difficult to see that

$$\hat f(\xi)\overline{\hat g(\xi+2\pi\mathsf{k})}=0\qquad\text{for all }\mathsf{k}\in[\Omega+\mathbb{Z}^d]\backslash\{\omega_0\}.$$

Now it is easy to see that (3.3) becomes

$$\delta(\omega_0)\int_{\mathbb{R}^d}\hat f(\xi)\overline{\hat g(\xi)}\overline{\hat\phi(\xi)}\hat\phi(\xi)d\xi$$
$$=\int_{\mathbb{R}^d}\hat f(\xi)\overline{\hat g(\xi+2\pi\omega_0)}\overline{\hat\phi(\xi)}\hat\phi(\xi+2\pi\omega_0)\cdot\sum_{\ell\in\{1\le n\le r\,:\,\omega\in\Omega_{\mathsf{M}_n}\}}\overline{\widehat{a_\ell}(\xi)}\widehat{a_\ell}(\xi+2\pi\omega_0)d\xi.$$

Note that $\hat\phi(\xi)\neq 0$ for almost every $\xi\in\mathbb{R}$. Since the above identity holds for all functions f,g in $L^2(\mathbb{R}^d)$ as long as the support of $\hat f$ is contained inside $B_\epsilon(\xi_0)$ and the support of $\hat g$ is contained inside $B_\epsilon(\xi+2\pi\omega_0)$, we now easily deduce from the above identity that $\sum_{\ell\in\{1\le n\le r\,:\,\omega\in\Omega_{\mathsf{M}_n}\}}\overline{\widehat{a_\ell}(\xi)}\widehat{a_\ell}(\xi+2\pi\omega_0)=\delta(\omega_0)$ for almost every $\xi\in B_\epsilon(\xi_0)$. Since ξ_0 can be arbitrary, we now conclude that (2.1) holds.

The theorem is proved. □

If $r=1$ and $\mathsf{M}_1=I_d$ in Theorem 3.1 (ii), then we must have $\widehat{a_1}=1$ (that is, $a_1=\delta$) and consequently for this particular case, $\mathrm{T}_{a_1,I_d}v=v$ and $\mathrm{S}_{a_1,I_d}v=v$, that is, the data is just copied.

3.2. Tight Frame Structure. For any positive integer J, we now construct a tight affine-like frame $\mathcal{AS}_J$ in $L^2(\mathbb{R}^d)$ corresponding to the decomposition and reconstruction algorithms, (FAD) (Figure 2.2) and (FAR) (Figure 2.3), respectively. Let a be a mask on $\mathbb{Z}^d$ and M_0 be a $d\times d$ dilation matrix, for example, $\mathsf{M}_0=2I_d$ and a is a tensor product mask. We assume that

$$\hat\phi(\xi):=\prod_{j=1}^{\infty}\hat a((\mathsf{M}_0^T)^{-j}\xi),\qquad \xi\in\mathbb{R}^d$$

is a well-defined function in $L^2(\mathbb{R}^d)$ and we also assume that there exist filters $b_1,\dots,b_r$ such that

$$\hat a(\xi)\overline{\hat a(\xi+2\pi\omega)}+\sum_{\ell=1}^{r}\widehat{b_\ell}(\xi)\overline{\widehat{b_\ell}(\xi+2\pi\omega)}=\delta(\omega),\qquad \omega\in\Omega_{\mathsf{M}_0}.$$

Based on these, we then define

$$\widehat{\psi^\ell}(\mathsf{M}_0^T\xi):=\widehat{b_\ell}(\xi)\hat\phi(\xi),\qquad \ell=1,\dots,r.$$

Then the system

$$\mathcal{WS}(\psi^1,\dots,\psi^r):=\{\psi^1_{\mathsf{M}_0^j;\mathsf{k}},\dots,\psi^r_{\mathsf{M}_0^j;\mathsf{k}}\ :\ j\in\mathbb{Z},\mathsf{k}\in\mathbb{Z}^d\}$$

forms a tight frame in $L_2(\mathbb{R}^d)$ by the Unitary Extension Principle of [42] (see also [34, 35, 43]).

Next, we use this standard tight frame system to derive a new tight affine-like frame $\mathcal{AS}_J$ in $L^2(\mathbb{R}^d)$ corresponding to the decomposition and reconstruction algorithms, (FAD) (Figure 2.2) and (FAR) (Figure 2.3), respectively.

First, denote

$$\psi^{(0,\dots,0)}=\phi.$$

Then, let $\beta\in\mathcal{L}_J^L\cup\bigcup_{j=1}^J\mathcal{L}_j^H$ with $\beta\in\mathcal{L}^L_{J,\beta^{(1)}}$, $\beta^{(1)}\in\mathcal{L}^L_{J-1}$, or $\beta\in\mathcal{L}^H_{j,\beta^{(1)}}$, $\beta^{(1)}\in\mathcal{L}^L_{j-1}$ for some $j\in\{1,\dots,J\}$. Then we recursively define affine functions by

$$\widehat{\psi^\beta}(\mathsf{M}_\beta^T\xi):=\widehat{a_\beta}(\xi)\widehat{\psi^{\beta^{(1)}}}(\xi).$$

Further, let $\beta^{(2)} \in \mathcal{L}^L_{j-2,\mu^{(2)}}$, etc. Using this sequence, we recursively define matrices by

$$\mathsf{N}_\beta := \mathsf{M}_\beta^{-1}\mathsf{M}_{\beta^{(1)}}^{-1} \cdots \mathsf{M}_{\beta^{(j-1)}}^{-1}\mathsf{M}_{\beta^{(j)}}^{-1}.$$

We now define the associated affine-like system in the following way:

DEFINITION 3.2. *Retaining the introduced notions and definitions, the* affine-like system $\mathcal{AS}_J$ *is defined by*

$$\mathcal{AS}_J := \{\psi^1_{\mathsf{M}_0^j;\mathsf{k}}, \ldots, \psi^r_{\mathsf{M}_0^j;\mathsf{k}} \ : \ j \geq J, \mathsf{k} \in \mathbb{Z}^d\} \cup \{\psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{k}} \ : \ \mathsf{k} \in \mathbb{Z}^d, \beta \in \mathcal{L}^L_J \cup \bigcup_{j=1}^J \mathcal{L}^H_j\},$$

where we also employ the notation introduced in (1.1).

We next show that this system – although in general not forming an orthonormal basis – still always constitutes a tight frame.

THEOREM 3.3. *For any positive integer J, $\mathcal{AS}_J$ is a tight frame for $L^2(\mathbb{R}^d)$.*

Proof. By Theorem 3.1, it is not difficult to deduce that

$$\sum_{\beta \in \mathcal{L}^L_j \cup \mathcal{L}^H_j} \sum_{\mathsf{k} \in \mathbb{Z}^d} |\langle f, \psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{k}}\rangle|^2 = \sum_{\beta \in \mathcal{L}^L_{j-1}} \sum_{\mathsf{k} \in \mathbb{Z}^d} |\langle f, \psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{k}}\rangle|^2.$$

Now from the above relation, we see that

$$\sum_{\beta \in \mathcal{L}^L_J \cup \bigcup_{j=1}^J \mathcal{L}^H_j} \sum_{\mathsf{k} \in \mathbb{Z}^d} |\langle f, \psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{k}}\rangle|^2 = \sum_{\mathsf{k} \in \mathbb{Z}^d} |\langle f, \psi^{(0,\ldots,0)}_{\mathsf{M}_0^J;\mathsf{k}}\rangle|^2 = \sum_{\mathsf{k} \in \mathbb{Z}^d} |\langle f, \phi_{\mathsf{M}_0^J;k}\rangle|^2.$$

By our given assumption on ϕ and $\psi^1, \ldots, \psi^r$, it is known that $\{\phi_{\mathsf{M}_0^J;\mathsf{k}} \ : \ \mathsf{k} \in \mathbb{Z}^d\} \cup \{\psi^\ell_{\mathsf{M}_0^j;\mathsf{k}} \ : \ j \geq J, \mathsf{k} \in \mathbb{Z}^d, \ell = 1, \ldots, r\}$ is a tight frame for $L^2(\mathbb{R}^d)$. Consequently,

$$\sum_{\beta \in \mathcal{L}^L_J \cup \bigcup_{j=1}^J \mathcal{L}^H_j} \sum_{\mathsf{k} \in \mathbb{Z}^d} |\langle f, \psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{k}}\rangle|^2 + \sum_{j=J}^\infty \sum_{\ell=1}^r \sum_{\mathsf{k} \in \mathbb{Z}^d} |\langle f, \psi^\ell_{\mathsf{M}_0^j;\mathsf{k}}\rangle|^2 = \|f\|^2_{L^2(\mathbb{R}^d)}.$$

Hence, $\mathcal{AS}_J$ is a tight frame for $L^2(\mathbb{R}^d)$. ☐

Next we analyze the approximation order of the affine-like system $\mathcal{AS}_J$. For this, for $\tau \geq 0$, we recall that $H^\tau(\mathbb{R}^d)$ consists of all functions $f \in L^2(\mathbb{R}^d)$ satisfying

$$|f|^2_{H^\tau(\mathbb{R}^d)} := \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} |\hat{f}(\xi)|^2 \|\xi\|^{2\nu} d\xi < \infty.$$

We say that a mask $a : \mathbb{Z}^d \mapsto \mathbb{C}$ has τ sum rules if

$$\sum_{\mathsf{k} \in \mathbb{Z}^d} a(\mathsf{n} + \mathsf{M}_0\mathsf{k})(\mathsf{n} + \mathsf{M}_0\mathsf{k})^\beta = \sum_{\mathsf{k} \in \mathbb{Z}^d} a(\mathsf{M}_0\mathsf{k})(\mathsf{M}_0\mathsf{k})^\beta$$

for all $\mathsf{n} \in \mathbb{Z}^d$ and for all $\beta = (\beta_1, \ldots, \beta_d) \in (\mathbb{N} \cup \{0\})^d$ such that $0 \leq \beta_1, \ldots, \beta_d < \tau, \beta_1 + \cdots + \beta_d < \tau$.

It follows from Theorem 3.3 that for any $f \in L^2(\mathbb{R}^d)$, expanding f under the tight frame $\mathcal{AS}_J$, we have

$$f = \sum_{\beta \in \mathcal{L}^L_J \cup \bigcup_{j=1}^J \mathcal{L}^H_j} \sum_{\mathsf{k} \in \mathbb{Z}^d} \langle f, \psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{k}}\rangle \psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{k}} + \sum_{j=J}^\infty \sum_{\ell=1}^r \sum_{\mathsf{k} \in \mathbb{Z}^d} \langle f, \psi^\ell_{\mathsf{M}_0^j;\mathsf{k}}\rangle \psi^\ell_{\mathsf{M}_0^j;\mathsf{k}}$$

with the series converging absolutely in $L^2(\mathbb{R}^d)$. We now consider the truncated series

$$P_J f := \sum_{\beta \in \mathcal{L}_J^L \cup \bigcup_{j=1}^J \mathcal{L}_j^H} \sum_{\mathsf{k} \in \mathbb{Z}^d} \langle f, \psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{k}} \rangle \psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{k}}.$$

By the same argument as in Theorem 3.3, it is not difficult to deduce that

$$P_J f = \sum_{\mathsf{k} \in \mathbb{Z}^d} \langle f, \phi_{\mathsf{M}_0^J;\mathsf{k}} \rangle \phi_{\mathsf{M}_0^J;\mathsf{k}}$$

The next theorem now follows from well-known results (see, e.g., [18, 28, 30, 32, 36] and references therein).

THEOREM 3.4. *Let $\mathcal{AS}_J$ be the tight frame for $L^2(\mathbb{R}^d)$ introduced in Definition 3.2. Suppose that M_0 is isotropic, that is, M_0 is similar to a diagonal matrix with all entries having the same modulus. If the mask a has τ sum rules, then $\mathcal{AS}_J$ has approximation order τ, that is, there exists a positive constant C such that*

$$\|f - P_J f\|_{L^2(\mathbb{R}^d)} \le C |\det \mathsf{M}_0|^{-\tau J/d} |f|_{H^\tau(\mathbb{R}^d)} \qquad \text{for all } f \in H^\tau(\mathbb{R}^d), J \in \mathbb{N}.$$

4. Shearlet Systems With an Associated AMRA. In this section, we will apply our general framework to the special case of shearlets. This leads to shearlet systems, especially compactly supported ones, associated with an AMRA structure and fast decomposition and reconstruction algorithms. We anticipate that our considerations will improve the applicability of shearlets to various applications, in particular, frame based image restorations.

We first present a very general construction, which – as shown in the second subsection – can be utilized to construct various shearlet systems with an associated AMRA and fast decomposition and reconstruction algorithms. Subsection 4.3 is then concerned with 3D shearlet constructions.

4.1. General Construction of an AMRA. We first follow up on the comments after Corollary 2.2, in which it was pointed out that the construction of a sequence of dilation matrices M_ℓ, $1 \le \ell \le r$, i.e., invertible integer matrices and filters a_ℓ, $\ell = 1, \dots, r$ satisfying (2.1), depends only on the generated lattice $\mathsf{M}_\ell \mathbb{Z}^d$, $1 \le \ell \le r$. Hence we first present a very general construction of such a sequence of matrices and filters provided that the matrices satisfy $\mathsf{M}_\ell \mathbb{Z}^d = \mathsf{M}\mathbb{Z}^d$ for all $1 \le \ell \le r$ for some matrix M, which was the hypothesis of Corollary 2.2.

For this, we first fix a sublattice Γ of $\mathbb{Z}^d$. Obviously, there exist many $d \times d$ matrices M such that $\mathsf{M}\mathbb{Z}^d = \Gamma$. We now construct a set of tight frame filters for such a lattice Γ so that

$$\sum_{\ell=1}^r \widehat{a_\ell}(\xi) \overline{\widehat{a_\ell}(\xi + 2\pi\omega)} = \delta(\omega), \qquad \omega \in \Omega_{\mathsf{M}}. \tag{4.1}$$

As already mentioned before, the above equations in (4.1) only depend on Ω_{M}, which in turn only depend on the lattice $\mathsf{M}\mathbb{Z}^d = \Gamma$. As showed in [29, Corollary 3.4], there exist two integer matrices E and F such that $|\det E| = |\det F| = 1$ and

$$\mathsf{M} = EDF, \qquad D = \operatorname{diag}(\mathsf{d}_1, \dots, \mathsf{d}_m, 1, \dots, 1), \quad \mathsf{d}_1 \ge \dots \mathsf{d}_m > 1.$$

For the dilation matrix $\mathsf{N} := \mathrm{diag}(\mathsf{d}_1, \ldots, \mathsf{d}_m)$, we first construct a tensor product tight affine frame filter bank as follows. For each $\mathsf{d}_n > 1$, one can easily construct a one-dimensional tight affine frame filter bank $u_\ell, \ell = 1, \ldots, r_{\mathsf{d}_n}$ such that

$$\sum_{\ell=1}^{r_{\mathsf{d}_n}} \widehat{u_\ell}(\xi)\overline{\widehat{u_\ell}(\xi + 2\pi\omega)} = \delta(\omega), \qquad \omega \in \{0, \tfrac{1}{\mathsf{d}_n}, \ldots, \tfrac{\mathsf{d}_n - 1}{\mathsf{d}_n}\}. \tag{4.2}$$

For the construction of (compactly supported) one-dimensional tight affine frame filters, see [18, 27, 31, 32, 42].

Now we construct a tensor product filter bank by

$$U_{(\ell_1,\ldots,\ell_m)}(\beta_1, \ldots, \beta_d) := u_{\ell_1}(\beta_1) \cdots u_{\ell_m}(\beta_m), \qquad \beta_1, \ldots, \beta_d \in \mathbb{Z}.$$

This generates a total of $r := r_{\mathsf{d}_1} \cdot \ldots \cdot r_{\mathsf{d}_m}$ filters. We reorder them as $U_1, \ldots, U_r$. By (4.2), one can easily check that

$$\sum_{\omega \in \Omega_{\mathsf{M}}} \widehat{U_\ell}(\xi)\overline{\widehat{U_\ell}(\xi + 2\pi\omega)} = \delta(\omega), \qquad \xi \in \mathbb{R}^d.$$

Now we define $\widehat{a_\ell}(\xi) := \widehat{U_\ell}(E^T\xi)$, $\xi \in \mathbb{R}^d, \ell = 1, \ldots, r$. Then we have

$$\sum_{\omega \in \Omega_{\mathsf{M}}} \widehat{a_\ell}(\xi)\overline{\widehat{a_\ell}(\xi + 2\pi\omega)} = \delta(\omega), \qquad \xi \in \mathbb{R}^d. \tag{4.3}$$

In other words, using a tensor product construction, for any sublattice of $\mathbb{Z}^d$ generated by $\mathsf{M}\mathbb{Z}^d$, we can easily obtain a set of (compactly supported) tight affine(-like) frame filters satisfying (4.3). It is very important to notice that the above construction only depends on the lattice $\mathsf{M}\mathbb{Z}^d$ instead of the dilation matrix M itself. For more details on construction of high-dimensional wavelet filters using the tensor product methods, see [29, 31].

Next, we shall present a general construction of tight affine(-like) frame filters to fulfill the need in this paper. Suppose that we are given a group of dilation matrices $\mathsf{M}_\ell, \ell = 1, \ldots, r$. To achieve directionality, as discussed above, only the dilation matrices for the low-pass filters will be important. In other words, if shear matrices are involved, i.e., direction-based matrices, they play an important role for low-pass filter only and for high-pass filters, the choice of the dilation matrices does not matter, since no further decomposition will be performed for high-pass coefficients.

We group these dilation matrices into subgroups according to their lattices $\mathsf{M}_\ell\mathbb{Z}^d$: if the lattices $\mathsf{M}_\ell = \mathsf{M}_{\ell'}$ are the same, then $\mathsf{M}_\ell, \mathsf{M}_{\ell'}$ are grouped into the same group. For each dilation matrix in a subgroup, we only use a fixed set of tight affine frame filters that are constructed for that lattice. In other words, for a given lattice, we have a set of tight affine frame filters and we have complete freedom in choosing the direction-based matrices, for instance, shear matrices, to achieve directionality as long as the resulting lattice is the same given lattice.

Suppose now that we choose N sets of such tight affine frame filters for all the groups of dilation matrices. Since these N sets of tight affine frames are completely independent, when we put them together to get one whole set of tight affine frame filters, we have to multiply the factor $\frac{1}{\sqrt{N}}$ to every involved filter in the set. Now it is straightforward to see that the total collection of all such N sets of renormalized tight affine frame filters indeed forms a collection of tight affine frame filters satisfying

(2.1), where r is the total number of all the involved filters. It is also very important to notice that the renormalization of the filters does not reduce the directionality of the tight affine frame system, since the coefficients in the same band has the same ordering of magnitude as the one without renormalization.

4.2. Shearlet Constructions. We next focus on explicit shearlet constructions. For this, we recall that the dilation for shearlets is composed of a shear matrix S_s and a parabolic scaling matrix A_c (cf. Subsection 1.3). Thus, at each level of the decomposition, the matrices M_ℓ will be chosen as

$$\mathsf{M}_\ell = S_\ell A_4, \tag{4.4}$$

for some shear matrices S_ℓ whose selection should be driven by the particular application at hand, and A_4 is the parabolic scaling matrix, where the choice of the value 4 (in contrast to 2) avoids technicalities caused by square roots. As already elaborated upon before, which of those matrices will be labeled low- and which one high-pass is left to the user.

4.2.1. Shearlet Unitary Extension Principle and Fast Shearlet Transform. The requirements on the filters associated with general matrices of the type described in (4.4) at each level were already derived in Theorem 1.1, which we coined the 'Shearlet Unitary Extension Principle'.

Aiming towards examples for possible filters, we first observe that there exists only two different lattices in Theorem 1.1: Consider the product $S_\ell A_4$ with $S_\ell = \begin{pmatrix} 1 & s_\ell \\ 0 & 1 \end{pmatrix}$. If s_ℓ is an even integer, then $S_\ell A_4 \mathbb{Z}^2 = A_4 \mathbb{Z}^2$, if it is an odd integer, then $S_\ell A_4 \mathbb{Z}^2 = S A_4 \mathbb{Z}^2$ with $S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Hence we only need to design two sets of tight shearlet frame filters in advance for the lattices $A_4\mathbb{Z}^2$ and $SA_4\mathbb{Z}^2$, respectively, following the general construction in Subsection 4.1. Then we have complete freedom in choosing the shear matrices S_ℓ to obtain a whole set of tight shearlet frame filters satisfying perfect reconstruction (see condition (i) in Theorem 1.1). In particular, notice that we can choose the filters such that the designed tight shearlet frame is compactly supported, for example, take $\mathsf{M}_0 = 2I_2$ and choose the tensor product 1D tight wavelet frames derived from the linear hat function as constructed in [42].

Each such a choice of filters for each level then leads to a Fast Adaptive Shearlet Decomposition associated with a Fast Adaptive Shearlet Reconstruction. Those two algorithms are described in Figures 4.1 and 4.2. Notice that here – as already in the general version of (FAD) and (FAR) – the shear matrices can be chosen differently at each level of the decomposition. We can envision that this adaption can be made flexible dependent on a quick analysis, thresholding, say, of the data outputted in the previous decomposition step. The algorithm leaves all those possibilities open. The great flexibility provided here should be utilizable for various applications including, in particular, frame based image restorations.

4.3. 3D Shearlets. Various applications such as the problem of geometric separation in image processing, for example, in astronomy when images of galaxies require separated analyzes of stars, filaments, and sheets call for directional representation systems for 3D data (see, e.g., [20, 21]). In this subsection, we will show that our framework conveys the potential to generate (compactly supported) shearlets for 3D signals alongside a fast flexible decomposition strategy.

FASD (FAST ADAPTIVE SHEARLET DECOMPOSITION)

Parameters:
- Data $v \in l(\mathbb{Z}^2)$.
- Number of iterations $J \in \mathbb{N}$.
- Set of all indices $\mathcal{L} := \bigcup_{j=1}^{J} \mathcal{L}_j^L \cup \bigcup_{j=1}^{J} \mathcal{L}_j^H$.
- Number of decompositions r_β for each position $\beta \in \mathcal{L}_{j-1}^L$, $j = 1, \ldots, J$.
- Separators between low- and high-pass parts $s_\beta \leq r_\beta$ for each position $\beta \in \mathcal{L}_{j-1}^L$, $j = 1, \ldots, J$.
- Sequence of finitely supported masks $a_\beta \in l(\mathbb{Z}^2)$, $\beta \in \mathcal{L}$, satisfying the conditions in Theorem 1.1.
- Sequence of shear matrices $(S_\beta)_{\beta \in \mathcal{L}}$ satisfying the conditions in Theorem 1.1.

Algorithm:

Set $v_{\mathbf{0}} = v_{(0,\ldots,0)} = v$.
For $j = 1, \ldots, J$ do
  For $\beta \in \mathcal{L}_{j,\alpha}^L$, $\alpha \in \mathcal{L}_{j-1}^L$ do
    Set $v_\beta = \mathrm{T}_{a_\beta, S_\beta A_4} v_\alpha$.
  end;
  For $\beta \in \mathcal{L}_{j,\alpha}^H$, $\alpha \in \mathcal{L}_{j-1}^L$ do
    Set $w_\beta = \mathrm{T}_{a_\beta, S_\beta A_4} v_\alpha$.
  end;
end;

Output:
- Low-pass coefficients v_β, where $\beta \in \mathcal{L}_J^L$.
- High-pass coefficients w_β, where $\beta \in \bigcup_{j=1}^{J} \mathcal{L}_j^H$.

FIG. 4.1. *The* FASD *Algorithm for a fast adaptive decomposition using shearlet systems.*

In fact, our framework also provides shearlets associated with an AMRA structure in higher dimensions. However, for the purpose of clarity we will focus on dimension 3. One possibility to extend the parabolic scaling is by defining

$$A_c = \begin{pmatrix} c & 0 & 0 \\ 0 & \sqrt{c} & 0 \\ 0 & 0 & \sqrt{c} \end{pmatrix}, \quad c > 0. \tag{4.5}$$

General shear matrices were for instance studied in [26]. But in order to present the idea for a 3D shearlet construction and don't cloud it with technicalities, the shear matrices introduced in [17] shall serve as an interesting example, and we define

$$S_{s_1,s_2} = \begin{pmatrix} 1 & s_1 & s_2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (s_1, s_2) \in \mathbb{R}^2. \tag{4.6}$$

We now set $\mathsf{M}_\ell = S_\ell A_4$, for some shear matrices S_ℓ of type (4.6), whose selection should be driven by the particular application at hand, and A_4 is the parabolic scaling matrix from (4.5).

Concerning a construction of filters, we can again employ the general construction from Subsection 4.1. By calculation, it is easy to see that we have only four different

FASR (FAST ADAPTIVE SHEARLET RECONSTRUCTION)

Parameters:
- Number of iterations $J \in \mathbb{N}$.
- Set of all indices $\mathcal{L} := \bigcup_{j=1}^{J} \mathcal{L}_j^L \cup \bigcup_{j=1}^{J} \mathcal{L}_j^H$.
- Number of reconstructions r_β for each position $\beta \in \mathcal{L}_{j-1}^L$, $j = 1, \ldots, J$.
- Separators between low- and high-pass parts $s_\beta \leq r_\beta$ for each position $\beta \in \mathcal{L}_{j-1}^L$, $j = 1, \ldots, J$.
- Low-pass coefficients v_β, where $\beta \in \mathcal{L}_J^L$.
- High-pass coefficients w_β, where $\beta \in \bigcup_{j=1}^{J} \mathcal{L}_j^H$.
- Sequence of finitely supported masks $a_\beta \in l(\mathbb{Z}^2)$, $\beta \in \mathcal{L}$, satisfying the conditions in Theorem 1.1.
- Sequence of shear matrices $(S_\beta)_{\beta \in \mathcal{L}}$ satisfying the conditions in Theorem 1.1.

Recursive Algorithm:

1) For $j = J-1, \ldots, 0$ do
2) For $\alpha \in \mathcal{L}_j^L$ do
3) Set $v_\alpha = \sum_{\beta \in \mathcal{L}_{j+1,\alpha}^L} \mathrm{S}_{a_\beta, S_\beta A_4} v_\beta + \sum_{\gamma \in \mathcal{L}_{j+1,\alpha}^H} \mathrm{S}_{a_\gamma, S_\beta A_4} w_\gamma$.
4) end;
5) end;

Output:
- Data $v_{(0,\ldots,0)} \in l(\mathbb{Z}^d)$.

FIG. 4.2. *The* FASR *Algorithm for a fast adaptive reconstruction using shearlet systems.*

sets of sublattices, according to even or odd s_1 and s_2. Using tensor products, we can design four sets of tight affine frame filters in advances. Then we have the complete freedom in choosing the shear matrices S_{s_1,s_2} to obtain a whole set of tight frame filters.

We would like to mention that although focussing on 3D in this subsection, our framework can also be used to generate shearlet systems in higher dimensions using the general definition of shear matrices in [26]. The parabolic scaling can be extended to higher dimension, for instance, as it was done in [17], but various other choices are possible, which will be explored in future work.

5. Further Examples of Directional Systems associated with an AMRA.

To show the broad applicability of the AMRA-structure, in this section, we will present further examples of systems which have the potential to analyze the directional content of images and signals.

5.1. Shearlet-Like Construction.

The resolution of anisotropic structures requires paramatrization of directions. The most natural way seems to be the choice of parametrization through angles. However, angles are related with the rotation operator, which is incompatible with digital lattices. This observation led to the introduction of shearlets, which measure directionality by slope rather than angle thereby greatly supporting the treating of the digital setting. However, it seems natural to seek for more general operators which leave lattice structures invariant while being somewhat close to rotation. These might differ from the shear matrices utilized in the

shearlet definition, and provide an improvement on uniform treatment of all angles. We would like to remind the reader that in the shearlet construction (cf. Subsection 1.3) the partition into cones serves as a means to allow a more uniform treatment of the angles, since the 'pure' shearlet system generated by the shearlet group is strongly biased towards one axis (for more details we refer the interested reader to [24]).

We can model this optimization problem in the following way:

$$(\mathrm{R{\scriptsize OT}}_\theta) \quad B_\theta = \ \operatorname{argmin}_{B\in M(\mathbb{Z},2)} \int_0^{2\pi} \|(R_\theta - B)\begin{pmatrix}\cos t\\ \sin t\end{pmatrix}\|_2^2\,dt \quad \text{s.t. } \det(B)=1,$$

where R_θ, $\theta \in [0,2\pi)$ is planar rotation by θ radians.

This optimization problem can be explicitly solved, which is done in the following result:

THEOREM 5.1. *For $\theta \in [0,2\pi)$, we denote the solution of $(\mathrm{R{\scriptsize OT}}_\theta)$ by B_θ. Then*

$$B_\theta = \begin{cases} I & : \ 2\pi - \arcsin(\frac{1}{2}) \le \theta < \arcsin(\frac{1}{2}),\\ S_{-1} \ \mathit{or} \ S_1^T & : \ \arcsin(\frac{1}{2}) \le \theta < \frac{\pi}{4},\\ S_1^{(1)} \ \mathit{or} \ S_1^{(2)} & : \ \frac{\pi}{4} < \theta \le \arccos(\frac{1}{2}),\\ S_0^{(1)} & : \ \arccos(\frac{1}{2}) < \theta \le \pi - \arccos(\frac{1}{2}),\\ S_{-1}^{(1)} \ \mathit{or} \ S_{-1}^{(2)} & : \ \pi - \arccos(\frac{1}{2}) < \theta \le \frac{3\pi}{4},\\ -S_1 \ \mathit{or} \ -S_{-1}^T & : \ \frac{3\pi}{4} < \theta \le \pi - \arcsin(\frac{1}{2}). \end{cases}$$

The remaining six cases can be determined by transposition of the previous matrices, in particular,

$$B_\theta = \begin{cases} -I & : \ \pi - \arcsin(\frac{1}{2}) < \theta \le \pi + \arcsin(\frac{1}{2}),\\ -S_1^T \ \mathit{or} \ -S_{-1} & : \ \pi + \arcsin(\frac{1}{2}) < \theta \le \frac{5\pi}{4},\\ (S_{-1}^{(1)})^T \ \mathit{or} \ (S_{-1}^{(2)})^T & : \ \frac{5\pi}{4} < \theta \le \pi + \arccos(\frac{1}{2}),\\ (S_0^{(1)})^T & : \ \pi + \arccos(\frac{1}{2}) < \theta \le \frac{3\pi}{2} + \arcsin(\frac{1}{2}),\\ (S_1^{(1)})^T \ \mathit{or} \ (S_1^{(2)})^T & : \ \frac{3\pi}{2} + \arcsin(\frac{1}{2}) < \theta \le \frac{7\pi}{4},\\ S_{-1}^T \ \mathit{or} \ S_1 & : \ \frac{7\pi}{4} \le \theta < 2\pi - \arcsin(\frac{1}{2}), \end{cases}$$

where

$$S_s^{(1)} = \begin{pmatrix} s & -1\\ 1 & 0\end{pmatrix} \quad \mathit{and} \quad S_s^{(2)} = \begin{pmatrix} 0 & -1\\ 1 & s\end{pmatrix}.$$

Proof. Set $B = (b_{ij})_{1\le i,j\le 2}$. First,

$$\begin{aligned} &\|(R_\theta - B)\begin{pmatrix}\cos t\\ \sin t\end{pmatrix}\|_2^2\\ &= [(\cos\theta - b_{11})^2 + (\sin\theta - b_{21})^2]\cos^2 t + [(-\sin\theta - b_{12})^2 + (\cos\theta - b_{22})^2]\sin^2 t\\ &\quad +2[(\cos\theta - b_{11})(-\sin\theta - b_{12}) + (\sin\theta - b_{21})(\cos\theta - b_{22})]\cos t\,\sin t. \end{aligned}$$

Hence

$$\begin{aligned} &\int_0^{2\pi} \|(R_\theta - B)\begin{pmatrix}\cos t\\ \sin t\end{pmatrix}\|_2^2\,dt\\ &= \pi[(\cos\theta - b_{11})^2 + (-\sin\theta - b_{12})^2 + (\sin\theta - b_{21})^2 + (\cos\theta - b_{22})^2]. \end{aligned}$$

We now aim to minimize the term on the RHS over $b_{11}, b_{12}, b_{21}, b_{22} \in \mathbb{Z}$ under the constraint that $\det B = 1$. Certainly, the solution to

$$\min_{B=(b_{ij})_{i,j}\in M(\mathbb{Z},2)} [(\cos\theta - b_{11})^2 + (-\sin\theta - b_{12})^2 + (\sin\theta - b_{21})^2 + (\cos\theta - b_{22})^2], \quad (5.1)$$

by denoting the elements in R_θ by r^θ_{ij}, is

$$b_{ij} = \mathrm{round}(r^\theta_{ij}) \quad \text{for all } i,j \in \{1,2\}.$$

where round$(\cdot)$ is the classical rounding to the nearest integer. Notice that $b_{ij} \in \{-1,0,1\}$. The next closest solution (possibly not unique) can be determined by choosing $i_0, j_0 \in \{1,2\}$ such that

$$\|\mathrm{round}(r^\theta_{i_0,j_0}) - r^\theta_{i_0,j_0}\|_2 \geq \|\mathrm{round}(r^\theta_{i,j}) - r^\theta_{i,j}\|_2 \quad \text{for all } i,j \in \{1,2\}.$$

Then choose

$$b_{i_0,j_0} = \begin{cases} \mathrm{sgn}(b_{i_0,j_0}) \cdot 1 & : \ \mathrm{round}(b_{i_0,j_0}) = 0, \\ 0 & : \ \mathrm{round}(b_{i_0,j_0}) \neq 0 \end{cases}$$

and

$$b_{ij} = \mathrm{round}(r^\theta_{i,j}) \quad \text{for all } (i,j) \neq (i_0,j_0).$$

In fact, we will see that provided the solution to (5.1) does not satisfies the constraint $\det B = 1$, then the just described closest one does, hence is a solution of (ROT$_\theta$).

We now consider the case $0 \leq \theta < \frac{\pi}{2}$. The other cases can be handled similarly.

- If $0 \leq \theta < \arcsin^{-1}(\frac{1}{2})$, then the closest solution is R_θ, which satisfies the constraint.

- If $\arcsin^{-1}(\frac{1}{2}) \leq \theta < \arccos^{-1}(\frac{1}{2})$, then the solution to (5.1) is

  $$A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

  for which $\det A = 2$. The closest one to this – as just described – is

  $$B = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad \text{for } \theta < \frac{\pi}{4},$$

  and

  $$B = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix} \text{ or } \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{for } \theta \geq \frac{\pi}{4}.$$

  These integer matrices obviously satisfy the constraint $\det B = 1$.

- If $\arccos^{-1}(\frac{1}{2}) \leq \theta < \frac{\pi}{2}$, then the closest solution is R_θ, which satisfies the constraint.

This proves the theorem. □

It is interesting to notice that all these matrices are in fact shear matrices or closely related to shear matrices.

This consideration now enables us to propose a shearlet-like system with presumable improved uniformity of the treatment of all angles. For this, we let θ_ℓ, $\ell = 1, \dots, 12$ be representatives of the angles giving rise to the twelve different cases discussed in Theorem 5.1. We choose

$$\mathsf{M}_\ell = B_{\theta_\ell} A_4, \quad \ell = 1, \dots, 12,$$

for each level leaving the 'cut' between low- and high-pass parts to the user. The sequence of associated filters a_ℓ, $\ell = 1, \dots, 12$ can be chosen similarly as in Section 4.

5.2. An Isotropic System for detecting Anisotropic Phenomena. The following is some explanation for the tight frame in Definition 3.2 about how we can achieve directionality by using the simplest system: dilation $\mathsf{M} = 2I_2$ and tensor product filters, but using shear matrices in the middle. It is very important to notice that for this case, the construction of all the filters in Corollary 2.2 (ii) is independent of the choice of the shear matrix. This simple fact has been noticed in [29] and makes the construction of multivariate wavelets very simple.

From now on, we assume that M is fixed (say, $\mathsf{M} = 2I_d$) and all filters are obtained by a tensor product construction. Therefore, these filters give preferences to horizontal and vertical direction. In other words, they are suitable for handling only edges having either horizontal or vertical orientation.

Let us look at the elements in $\mathcal{AS}_J$ introduced in Definition 3.2. For the functions in level 1, we have $\alpha \in \mathcal{L}_1^L \cup \mathcal{L}_1^H$, that is, $\alpha = (\ell, 0, \dots, 0)$ with $0 < \ell \le r_\alpha$. Then $\psi^\alpha_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{n}}$ are given by

$$\begin{aligned}\psi^\alpha(\mathsf{M}_\alpha^{-1}\mathsf{M}_0^J \cdot - \mathsf{n}) &= |\det(\mathsf{M}_\alpha)| \sum_{\mathsf{k}\in\mathbb{Z}^d} a_\alpha(\mathsf{k}) \phi(\mathsf{M}_0^J \cdot - \mathsf{M}_\alpha \mathsf{n} - \mathsf{k}) \\ &= |\det(\mathsf{M}_0)| (a_\alpha \star \phi)(\mathsf{M}_0^J \cdot - \mathsf{M}_\alpha \mathsf{n}).\end{aligned}$$

Note that $\psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{n}}$ with a high-pass filter has the same directionality as the filter a_β, that is, only either horizontal or vertical direction. Note that we assume to have horizontal or vertical directionality for high-pass filters only, while for low-pass filters, we assume that they are isotropic and there is no directionality for them. In other words, for low-pass filters, the filters have no directionality.

The directionality comes at the next level starting with level 2. For $\beta \in \mathcal{L}_2^L \in \mathcal{L}_2^H$, we have $\beta = \alpha + (0, \ell, 0, \dots, 0)$ with $\alpha \in \mathcal{L}_1^L$ and $0 < \ell \le r_\beta$. Note $\mathsf{N} = \mathsf{M}_\beta^{-1}\mathsf{M}_\alpha^{-1}$. So, $\psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{n}}$ are given by

$$\psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{n}} = \psi^\beta(\mathsf{M}_\beta^{-1}\mathsf{M}_\alpha^{-1}\mathsf{M}_0^J \cdot - \mathsf{n}) = |\det(\mathsf{M}_0)| \sum_{\mathsf{k}\in\mathbb{Z}^d} a_\beta(\mathsf{k}) \psi^\alpha(\mathsf{M}_\alpha^{-1}\mathsf{M}_0^J \cdot - \mathsf{M}_\beta \mathsf{n} - \mathsf{k}).$$

That is, we similarly have

$$\psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{n}} = |\det(\mathsf{M}_0)| (a_\beta \star \psi^\alpha)(\mathsf{M}_\alpha^{-1}\mathsf{M}_0^J \cdot - \mathsf{M}_\beta \mathsf{n}).$$

To see the directionality better, we rewrite it in terms of ϕ as follows:

$$\begin{aligned}\psi^\beta_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{n}} &= |\det(\mathsf{M}_0)| \sum_{\mathsf{k}\in\mathbb{Z}^d} a_\beta(\mathsf{k}) \psi^\alpha(\mathsf{M}_\alpha^{-1}\mathsf{M}_0^J \cdot - \mathsf{M}_\beta \mathsf{n} - \mathsf{k}) \\ &= |\det(\mathsf{M}_0)|^2 \sum_{\mathsf{k}\in\mathbb{Z}^d} a_\beta(\mathsf{k}) (a_\alpha \star \phi)(\mathsf{M}_0^J \cdot - \mathsf{M}_\alpha \mathsf{M}_\beta \mathsf{n} - \mathsf{M}_\alpha \mathsf{k}).\end{aligned}$$

To see the directionality of $\psi^{\beta}_{\mathsf{N}_\beta \mathsf{M}_0^J;\mathsf{n}}$, it suffices to see it for $\mathsf{n} = 0$, since all others are just shifts of $\mathsf{n} = 0$ (that is, they have the same directionality). In this case,

$$\psi^{\beta}_{\mathsf{N}_\beta \mathsf{M}_0^J;0} = |\det(\mathsf{M}_0)|^2 \sum_{\mathsf{k}\in\mathbb{Z}^d} a_\beta(\mathsf{k})(a_\alpha \star \phi)(\mathsf{M}_0^J \cdot - \mathsf{M}_\alpha \mathsf{k}).$$

Note that a_α is a low-pass filter and therefore, it has no directionality. In other words, the function $a_\alpha \star \phi$ has most of its energy concentrated about the origin. Though M_0^J can also yield directionality by rotating the function, let us ignore this possibility here and simply assume that $\mathsf{M}_0 = 2I_d$ (or even simply take $J = 0$).

Now let us argue that if a_β, as a high-pass filter, has directionality either in the horizontal or vertical direction, then the function $\psi^{\beta}_{\mathsf{N}_\beta \mathsf{M}_0^J;0}$ will have directionality other than horizontal and vertical, induced by $\mathsf{M}_\alpha = S_\alpha \mathsf{M}_0$, where S_α is some shear matrix. For simplicity, we assume a_β is a filter with horizontal directionality. Therefore, $a_\beta(\mathsf{k})$ concentrates most of its energy for k on/around the x-axis. But the energy of the term $(a_\alpha \star \phi)(\mathsf{M}_0^J \cdot - \mathsf{M}_\alpha \mathsf{k})$ concentrates around the point $\mathsf{M}_0^{-J}\mathsf{M}_\alpha \mathsf{k}$. Therefore, for $\mathsf{k} = (k, 0, \dots, 0)$ with $k \in \mathbb{Z}$ (this is the line where most energy of a_β is concentrating), the energy location for $\psi^{\beta}_{\mathsf{N}_\beta \mathsf{M}_0^J;0}$ in above sum is put at the location $\mathsf{M}_0^{-J}\mathsf{M}_\alpha \mathsf{k} = \mathsf{M}_0^{1-J} S_\alpha \mathsf{k}$ instead of $\mathsf{M}_0^{1-J}\mathsf{k}$. In other words, the energy of the function $\psi^{\beta}_{\mathsf{N}_\beta \mathsf{M}_0^J;0}$ is aligned around the straight line: the image of the x-axis under the mapping $\mathsf{M}_0^{1-J} S_\alpha$. Since S_α acts like a rotation, the directionality is controlled by this mapping; the horizontal and vertical lines are mapped into other directions. By using different S_α (such as the power of some shear matrices), as many directions as required by the application at hand can be created.

REFERENCES

[1] J.-F. Cai, R. Chan, L. Shen and Z. Shen, *Restoration of chopped and nodded images by framelets*, SIAM J. Sci. Comput.,**30** (2008), 1205–1227.

[2] J.-F. Cai, R. H. Chan, L. Shen and Z. Shen, *Simultaneously Inpainting in Image and Transformed Domains*, Numer. Math. **112** (2009), 509–533.

[3] J.-F. Cai, R. H. Chan, L. Shen and Z. Shen, *Convergence analysis of tight framelet approach for missing data recovery*, Adv. Comput. Math. **31** (2009), 87–113.

[4] J.-F. Cai, R. H. Chan and Z. Shen, *A framelet-based image inpainting algorithm*, Appl. Comput. Harmon. Anal. **24** (2008), 131–149.

[5] J.-F. Cai, H. Ji, C. Liu and Z. Shen, *Blind motion deblurring using multiple images*, J. Comput. Phys., to appear.

[6] J.-F. Cai, H. Ji, C. Liu and Z. Shen, *Blind motion deblurring from a single image using sparse approximation*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009.

[7] J.-F. Cai, S. Osher and Z. Shen, *Linearized Bregman iteration for frame based image deblurring*, SIAM J. Imaging Sci. **2** (2009), 226–252.

[8] J.-F. Cai, S. Osher and Z. Shen, *Split bregman methods and frame based image restoration*, Multiscale Model. Simul., to appear.

[9] E. J. Candès, L. Demanet, D. L. Donoho and L. Ying, *Fast discrete curvelet transforms,* Multiscale Model. Simul. **5** (2006), 861–899.

[10] E. J. Candès and D. L. Donoho, *New tight frames of curvelets and optimal representations of objects with $C2$ singularities*, Comm. Pure Appl. Math. **56** (2004), 219–266.

[11] E. J. Candès and D. L. Donoho, *Continuous curvelet transform: I. Resolution of the wavefront set*, Appl. Comput. Harmon. Anal. **19** (2005), 162–197.

[12] E. J. Candès and D. L. Donoho, *Continuous curvelet transform: II. Discretization of frames*, Appl. Comput. Harmon. Anal. **19** (2005), 198–222.

[13] A. Chai and Z. Shen, *Deconvolution: A wavelet frame approach*, Numer. Math. **106** (2007), 529–587.

[14] R. H. Chan, T. F. Chan, L. Shen, and Z. Shen, *Wavelet algorithms for high-resolution image reconstruction*, SIAM J. Sci. Comput. **24** (2003), 1408–1432 (electronic).
[15] R. H. Chan, S. D. Riemenschneider, L. Shen, and Z. Shen, *Tight frame: an efficient way for high-resolution image reconstruction*, Appl. Comput. Harmon. Anal. **17** (2004), 91–115.
[16] R. H. Chan, Z. Shen, and T. Xia, *A framelet algorithm for enhancing video stills*, Appl. Comput. Harmon. Anal. **23** (2007), 153–17.
[17] S. Dahlke, G. Steidl, and G. Teschke, *The Continuous Shearlet Transform in Arbitrary Space Dimensions*, J. Fourier Anal. Appl., to appear.
[18] I. Daubechies, B. Han, A. Ron, and Z. Shen, *Framelets: MRA-based constructions of wavelet frames*, Appl. Comput. Harmon. Anal. **14** (2003), 1–46.
[19] M. N. Do and M. Vetterli, *The contourlet transform: an efficient directional multiresolution image representation*, IEEE Trans. Image Process. **14** (2005), 2091–2106.
[20] D. L. Donoho and G. Kutyniok, *Microlocal Analysis of the Geometric Separation Problem - I*, preprint.
[21] D. L. Donoho and G. Kutyniok, *Microlocal Analysis of the Geometric Separation Problem - II*, preprint.
[22] D. L. Donoho and G. Kutyniok, *Sparsity Equivalence of Anisotropic Decompositions*, preprint.
[23] G. Easley, D. Labate, and W. Lim, *Sparse Directional Image Representations using the Discrete Shearlet Transform*, Appl. Comput. Harmon. Anal. **25** (2008), 25–46.
[24] K. Guo, G. Kutyniok, and D. Labate, *Sparse Multidimensional Representations using Anisotropic Dilation and Shear Operators*, Wavelets and Splines (Athens, GA, 2005), Nashboro Press, Nashville, TN (2006), 189-201.
[25] K. Guo and D. Labate, *Optimally Sparse Multidimensional Representation using Shearlets*, SIAM J. Math. Anal. **39** (2007), 298–318.
[26] K. Guo, D. Labate, W. Lim, G. Weiss, and E. Wilson, *Wavelets with Composite Dilations and their MRA Properties*, Appl. Comput. Harmon. Anal. **20** (2006), 231–249.
[27] B. Han, *On dual wavelet tight frames*, Appl. Comput. Harmon. Anal. **4** (1997), 380–413.
[28] B. Han, *Approximation properties and construction of Hermite interpolants and biorthogonal multiwavelets*, J. Approx, Theory **110** (2001), 18–53.
[29] B. Han, *Symmetry property and construction of wavelets with a general dilation matrix*, Linear Algebra Appl. **353** (2002), 207–225.
[30] B. Han, *Vector cascade algorithms and refinable function vectors in Sobolev spaces*, J. Approx. Theory **124** (2003), 44–88.
[31] B. Han, Compactly supported tight wavelet frames and orthonormal wavelets of exponential decay with a general dilation matrix, *J. Comput. Appl. Math.*, **155** (2003), 43–67.
[32] B. Han, *Dual multiwavelet frames with high balancing order and compact fast frame transform*, Appl. Comput. Harmon. Anal. **26** (2009), 14–42.
[33] B. Han, *Paris of frequency-based nonhomogeneous dual wavelet frames in the distribution space*, Appl. Comput. Harmon. Anal. under revision, (2009)
[34] B. Han and Z. Shen, *Compactly supported symmetric C^∞ wavelets with spectral approximation order*, SIAM J. Math. Anal. **40** (2008), 905–938.
[35] B. Han and Z. Shen, *Dual wavelet frames and Riesz bases in Sobolev spaces*, Constr. Approx. **29** (2009), 369–406.
[36] R. Q. Jia, *Approximation properties of multivariate wavelets*, Math. Comp. **67** (1998), 647–665.
[37] G. Kutyniok, M. Shahram, and D. L. Donoho, *Development of a Digital Shearlet Transform Based on Pseudo-Polar FFT*, Wavelets XIII (San Diego, CA, 2009), 74460B-1 - 74460B-13, SPIE Proc. **7446**, SPIE, Bellingham, WA, 2009.
[38] G. Kutyniok and D. Labate, *Resolution of the Wavefront Set using Continuous Shearlets*, Trans. Amer. Math. Soc. **361** (2009), 2719–2754.
[39] G. Kutyniok and W.-Q. Lim, *Compactly supported shearlets are optimally sparse*, preprint.
[40] G. Kutyniok and T. Sauer, *Adaptive directional subdivision schemes and Shearlet Multiresolution Analysis*, SIAM J. Math. Anal. **41** (2009), 1436–1471.
[41] W. Lim, *Discrete Shearlet Transform: New Multiscale Directional Image Representation* Proceedings of SAMPTA'09, Marseille 2009.
[42] A. Ron and Z. Shen, *Affine systems in $L^2(\mathbb{R}^d)$: the analysis of the analysis operator*, J. of Funct. Anal. **148** (1997), 408–447.
[43] A. Ron and Z. Shen, *Affine systems in $L^2(\mathbb{R}^d)$: dual systems*, J. Fourier Anal. Appl. **3** (1997), 617–637.